Brit. J. Phil. Sci. 39 (1988), 1–50 *Printed in Great Britain*

501.05
B 774

Quantum Propensiton Theory: A Testable Resolution of the Wave/Particle Dilemma*

NICHOLAS MAXWELL

1 *Problems of Orthodox Quantum Theory*
2 *Inadequacy of Earlier Attempts to Solve the Problem*
3 *Propensiton Solution to the Wave/Particle Problem*
4 *From Orthodox to Propensiton Quantum Theory*
5 *Virtual, Potential and Actual Particles*
6 *Particle Creation and Annihilation as the Quantum Condition for Probabilistic Events to Occur*
7 *Laws Governing Probabilistic Events*
8 *Experimental Success of OQT Ensures Experimental Success of PQT*
9 *Crucial Experiments*
10 *Why has Quantum Propensiton Theory been Overlooked?*
11 *Conclusion*
12 *Appendix*

27

1 PROBLEMS OF ORTHODOX QUANTUM THEORY

What sort of physical objects are electrons, protons, photons, atoms, molecules—the entities of the quantum world—in view of the contradictory wave and particle properties that these objects appear to possess? This deserves to be regarded as *the* fundamental problem concerning the nature of the quantum world. It is above all this problem that we must solve if we are to have an adequate understanding of the quantum domain.

Orthodox quantum theory (OQT) *evades* and does not *solve* this key problem. The creators of OQT—Heisenberg, Bohr, Born, Dirac and others—decided, in effect, that no consistent, fully micro realistic theory of quantum objects evolving and interacting in space and time could be developed which did justice to both wave and particle aspects of quantum phenomena. As a result, they developed OQT as a theory which is solely about the results of performing

---

* I am especially grateful to the following for helpful criticisms of earlier drafts of this paper: Robert Seymour, Ian Thompson, Basil Hiley, Euan Squires, Gordon Fleming, Chris Isham, Richard Healey and Frank Arntzenius. Final revisions were made at the Center for Philosophy of Science at Pittsburgh University: I am grateful to the Center for providing me with the time to do the work.

measurements on (prepared ensembles of) quantum systems. The state vector ψ of OQT cannot be interpreted as specifying the actual physical state of the individual quantum system in physical space and time, because there is no solution to the wave/particle dilemma; rather ψ is to be interpreted as containing probabilistic information about the results of measuring diverse quantum observables, such as position, momentum, energy, spin. It is not possible to specify the instantaneous physical state of the individual quantum system in terms of the values of its quantum observables at that instant, because quantum systems do not in general, according to OQT, possess precise values of observables in the absence of measurement.

OQT has met with extraordinary experimental success. This has led many physicists to believe that it does not matter in the least, as far as physics itself is concerned, that OQT does not solve the wave/particle problem. In this they are wrong. OQT suffers from the following seven severe defects *as a physical theory* just because it fails to solve the wave/particle problem.

(i) *OQT cannot be given a micro realistic interpretation: the theory has no definite, characteristic physical ontology.* Fundamental dynamical theories of classical physics—such as Newton's theory of gravitation (NT) and Maxwell's theory of electromagnetism (MT)—can be interpreted micro realistically, as specifying the manner in which definite kinds of physical objects evolve and interact in physical space and time irrespective of whether the objects are undergoing measurement. Thus NT can be interpreted to be about point-particles, possessing inertial mass and gravitational charge, and therefore, as a result, being surrounded by a spherically symmetrical, rigid, infinite gravitational force-field which falls off in proportion to the inverse of the square of the distance. Likewise, MT can be interpreted to be about the electromagnetic field, specifying the precise way in which this evolves in space and time.

QT is above all a theory which is about micro objects or systems, and which seeks to predict and explain macro phenomena in terms of micro phenomena. If QT is to achieve this, it is clearly important to develop QT as a theory which can be interpreted micro realistically, like classical theories, as being about micro systems evolving and interacting in space and time—a theory with its own definite, characteristic physical ontology. But OQT cannot be given such a micro realistic interpretation, just because OQT provides no solution to the wave/particle problem.

Granted that we hold, with Galileo, Faraday, Maxwell, Boltzmann, Planck, Einstein and many others that it is a basic task of physics to improve our (conjectural) knowledge and understanding of the universe as it is in reality, independent of observation and measurement, then the failure of OQT to solve the key mystery of the nature of quantum objects as they are in reality, independent of measurement must be judged to be a serious defect indeed. However, many physicists, and some philosophers (e.g. van Fraassen 1980) reject this realist task for physics, and hold instead that physics has the more modest aim of developing theories which merely predict more and more phenomena more and more accurately (without necessarily describing unobserved reality). Granted this more modest instrumentalist aim, OQT cannot be held to be defective solely because it fails to tell us what kind of objects electrons and protons really are.

What follows can be regarded as providing a powerful case for rejecting instrumentalism (including van Fraassen's 'constructive empiricism') and adopting realism as the basic aim for physics instead. For I shall argue that *even if realism is rejected*, and OQT is assessed from a purely instrumentalist standpoint, nevertheless OQT must be judged to suffer from the following six severe defects, (ii) to (vii), it only being possible to overcome these six defects by developing a fully micro realistic version of QT, satisfying the demands of realism. In other words, even if instrumentalists are not interested in realism at all, nevertheless they are forced to adopt realism to the extent that non-realistic versions of QT, like OQT, suffer from severe defects, from the instrumentalist standpoint, which only a realist version of QT can overcome. And the point is quite general: some of the defects which OQT suffers from through not being realistically interpreted are the kind of defects which must plague any fundamental physical theory not realistically interpreted.

Here then are six defects of OQT which all physicists must take seriously, whatever their philosophy of physics, whether instrumentalist or realist—six defects which arise however as a consequence of defect (i), as a consequence of the non-realism of OQT *due to the lack of a solution to the wave/particle problem.*

(ii) *OQT is a very severely ad hoc theory, in a surreptitious and rarely noticed way, as a result of failing to solve the wave/particle problem.* The purely quantum mechanical part of OQT is not *ad hoc*; but this part of OQT is devoid of physical content in that it can issue in no physical predictions at all, because it lacks its own consistent quantum ontology. No combination of initial conditions and dynamical equations, formulated in purely quantum mechanical terms, can predict any actual physical state of affairs. On its own, OQT can, at most, only issue in *conditional* or *counterfactual* predictions about what would be the outcome *if* a measurement were to be performed. In order to issue in unconditional predictions, OQT must call upon some additional theory, with its own consistent physical ontology, for a specification of the physical states of preparation and measurement devices. As Bohr always emphasized, only OQT *plus some part of classical physics for a description of measurement* has genuine physical predictive content (Bohr 1949; see also Landau and Lifschitz 1958, p. 3). Attempts to dispense with classical physics by describing measuring instruments quantum mechanically must fail because such a purely quantum mechanical description can in turn only issue in predictions about what *would* occur *if* a measurement *were* to be made by some additional measuring instrument which must itself be described in terms of classical physics. (Such attempts must fail for other reasons as well: the dynamical equations of OQT assert that quantum states evolve deterministically, pure states never being converted into mixed states; measurement, however, is in general a probabilistic interaction, and one which *does* convert pure states into mixed states. For further details see Wigner 1967, ch. 12, Fine 1970, Maxwell 1972b.)

It is thus only the purely quantum mechanical part of OQT plus (some part of) classical physics (OQT+CP) which has any physical content, and thus constitutes a physical theory. But this hybrid theory, OQT+CP, is appallingly, grotesquely *ad hoc*, due to the fact that it is made up of two conceptually incoherent parts.

In recent years attempts have been made to develop a version of quantum theory (QT) applicable to macro phenomena in a quasi classical manner, and thus capable of weaning OQT of its conceptual dependence on classical physics (see Hepp 1972; Machida and Namiki 1984). If some such macro quantum theory (MQT) proves to be technically feasible, it would become possible to regard the physical theory of QT as being OQT+MQT rather than OQT+CP. But this does not help much: OQT+MQT must be almost exactly as *ad hoc* as OQT+CP. Freeing OQT of its dependence on CP in this way can do little to reduce the *ad hoc* character of the physical theory.

It is of course true that in order to check up on the predictions of a classical theory such as Newtonian theory (NT), we often need to employ additional physical theories, as when optical theory is used to check up on predictions of NT applied to the solar system. This does not mean, however, that NT is *ad hoc* in the same way in which OQT is. The difference is simply this. Because we can interpret NT as having its own consistent physical ontology (of massive, gravitationally charged particles), NT (plus specification of initial conditions) does issue in quite definite physical predictions about actual physical states of affairs—the positions and velocities of planets at definite times, for example—in the absence of optical or other physical theories, for measurement. NT is a physical theory with physical content in its own right; OQT is not.

(iii) *Despite its immense empirical success, OQT is seriously defective from the standpoint of enabling us to explain and understand quantum phenomena.* There are at least three reasons for holding this to be the case. (a) A basic task of QT is to predict and explain complex macro phenomena in terms of elementary micro phenomena—so that macro phenomena can be explained and understood as the outcome of interactions between vast numbers of micro systems. But this OQT cannot do, because the theory lacks a consistent model for micro systems, a consistent micro ontology (point (i) above). OQT can only specify and describe states of micro systems relative to prior classical descriptions of macro systems—preparation and measurement devices. Description of micro states presupposes, as a matter of conceptual necessity, description of macro states. That which is to be explained must be presupposed! Hence OQT cannot conceivably, even in principle, explain macro phenomena as arising solely as a result of interactions between large numbers of micro systems. (Instrumentalists may, with some justice, hold that this argument simply presupposes realism: this is not true however of the following two arguments (b) and (c).) (b) QT has the task of explaining the (approximate) empirical success of classical physics from purely quantum mechanical postulates. But this, again, OQT cannot do. In any physical application, OQT must presuppose (some part of) classical physics for an account of preparation and measurement devices. Once again, just that which is to be explained must be presupposed. (c) In order to be explanatory, a theory must not be *ad hoc.* But we have seen that the theory which has physical content, OQT+CP (or OQT+MQT), is very seriously *ad hoc.* Therefore, OQT is very seriously non-explanatory.

(iv) *OQT, regarded as a physical theory, is unacceptably imprecise.* On the face of it, QT is a fundamentally probabilistic theory. According to OQT, probabilistic events occur if and only if measurements are made—or at least if and only if measuring-type interactions occur. If OQT is to be a precise theory, it must specify precise physical conditions for probabilistic events to occur. But this cannot be done in terms of the imprecise notion of measurement. Physical processes cannot be precisely subdivided into those that do, and those that do not, constitute measuring-type processes. Furthermore, specifying measurement in terms of conscious observation, the occurrence of a macro process, a classical process, or an irreversible process, does not help as these notions are all irredeemably imprecise as well (Maxwell 1972b). Employing some MQT of macro quantum phenomena, as envisaged by Hepp or Machida and Namiki, cannot help much either, as any such MQT will be applicable to a great number of quantum systems, and will thus be highly imprecise from an elementary standpoint. OQT+CP (or OQT+MQT) is thus severely imprecise, in an irredeemable way, and to an unacceptable extent.

(v) *OQT is a seriously ambiguous theory, in that it is ambiguous as to whether probabilistic events occur at all.* Granted that a quantum mechanically described system S (or ensemble of such systems) is measured by a classically described measuring instrument M, OQT makes in general a probabilistic prediction about the outcome. One might suppose from this that OQT asserts unambiguously that probabilistic events occur when measuring-type interactions take place. But this is not correct. In principle the *deterministic* dynamical equations of OQT could be applied to the joint system S+M, in which case OQT predicts that S+M evolves deterministically until a further measurement is performed by an additional measuring instrument M*. This has led some to conclude that OQT is fundamentally a *deterministic* theory, probabilistic predictions emerging only because measuring instruments are in different quantum mechanical states when different particles are measured. Something like this must be assumed by all those who try to solve the so-called quantum 'problem of measurement' by trying to show that all measurement interactions evolve in accordance with the deterministic dynamical equations of OQT. A solution to this problem, conceived of in this way, would demonstrate the fundamentally *deterministic* character of OQT. In brief, OQT is only a fundamentally probabilistic theory in a highly *ambiguous* fashion.

(vi) *OQT is seriously restricted in scope.* It is standard practice these days to apply QT to states of the cosmos soon after the big bang, in physical conditions which preclude the very possibility of the existence of anything remotely corresponding to preparation and measurement devices. OQT cannot be applied in this way. Only a version of QT which has its own micro ontology could be thus applied.

Current theorizing about early states of the universe makes it desirable to be able to apply QT to the cosmos as a whole (thus creating the new discipline of quantum cosmology). Once again, OQT cannot be employed in this way, it being conceptually impossible that the cosmos as a whole should be subject to preparation and measurement!

(vii) *OQT cannot be generalized to include gravity.* Within the framework of OQT, a physical system only has a quantum state insofar as it is subject to preparation and measurement devices which are external, or additional, to the system in question. In order to quantize general relativity, space-time itself would need to be given quantum states. In order to do this within the framework of OQT, it would be necessary to postulate preparation and measurement devices external to space-time. No such devices can exist. Hence general relativity cannot be quantized within the framework of OQT.

In the light of the above seven defects, OQT must be declared to be a seriously unsatisfactory theory. (See also Maxwell 1972b, 1973, 1976a, 1982.) Even instrumentalists must reach this conclusion, since only defect (i) presupposes realism, whereas defects (ii) to (vii) do not. In fact, as I have already remarked, the above arguments do not just tell against an instrumentalist defence of OQT; they also tell against instrumentalism itself. For the six defects (ii) to (vii) all arise from defect (i)—from the failure of OQT to be open to a realist interpretation in the sense that the theory has its own consistent (possible) quantum ontology entirely independent of the ontology of classical physics. Furthermore, *any* fundamental physical theory which is not interpreted realistically as having its own (possible) ontology must inevitably suffer from defects (ii) and (iii)—and probably defects (vi) and (vii) as well: there is here, then, a general argument against instrumentalism, against 'constructive empiricism' (van Fraassen 1980), and *for* realism.

There is also the following additional argument which powerfully reinforces the arguments (i) to (vii) designed to show that OQT is a seriously defective physical theory, despite its immense empirical success. Elsewhere, I have shown that the widely held thesis that scientific theories are, in the end, to be judged solely with respect to empirical success and failure is untenable. Two kinds of criteria must always be employed in judging scientific theories: (1) *empirical* criteria, and (2) *non-empirical* criteria that have to do with the extent to which the theory is explanatory, non-*ad hoc*, unified, conceptually coherent, capable of fitting coherently into the best overall scientific understanding of the universe. Science requires both kinds of criteria equally. Without the non-empirical critera (2), the whole scientific enterprise breaks down: science would become overwhelmed by infinitely many empirically highly successful but grotesquely *ad hoc* theories, and all scientific knowledge, at the level of theory, becomes impossible. (For a detailed presentation and development of this argument see Maxwell, 1972a, 1974, 1976b, 1977, 1979, 1980, and especially 1984, Ch. 9.)

The attitude that we must adopt, then, in science if we are to be honest, is the following. In order to be scientifically acceptable, a theory must satisfy *both* criteria, (1) *and* (2) equally. A theory which satisfies beautifully non-empirical criteria (2) but fails dismally to satisfy empirical criteria (1), cannot be held to be a part of scientific knowledge, and must be rejected. But equally, *a theory which satisfies beautifully empirical criteria (1), but fails dismally to satisfy non-empirical criteria (2), cannot be held to be a part of scientific knowledge either, and must be rejected.*

This latter is the situation as far as OQT is concerned. The above seven points show, dramatically and decisively, that OQT is very seriously defective from the standpoint of non-empirical criteria (2)—from the standpoint, that is, of the search for explanation and understanding. (This is true even of point (vi), which concerns the inadequacy of OQT from the standpoint of attempting to understand early states of the cosmos, and the cosmos as a whole.)

The conclusion we ought to draw, then, is this. Despite its immense empirical success, QT given its orthodox interpretation cannot be held to be a part of scientific knowledge, and deserves to be rejected. We urgently need a better version of QT.

It is of decisive importance to appreciate that the above seven defects of OQT (including the so-called measurement problem) all arise because OQT *evades* and does not *solve* the wave/particle problem. For it is this *evasion* which makes it necessary to build the notions of observable and measurement into the orthodox concept of quantum state—thus creating the problems discussed above. *Solve* the wave/particle problem, and all this becomes unnecessary. It becomes possible to formulate QT as a (testable) theory about quantum objects *per se* evolving in space and time, the theory thus making no reference to observables, measurement or classical physics whatsoever. Measurement becomes a conceptually unproblematic physical process just like any other physical process, namely: quantum objects evolving in space and time in accordance with the laws of QT. The above seven defects vanish at a stroke. In short, in order to develop an acceptable version of QT, free of the above defects, free of the so-called measurement problem, the key problem that must be solved is the wave/particle problem—the problem of specifying a consistent ontology for the quantum domain. Einstein was absolutely correct when he remarked '. . . one simply cannot get around the assumption of reality—if only one is honest. Most . . . [physicists] simply do not see what sort of risky game they are playing with reality—reality as something independent of what is experimentally established.' (Einstein, 1950.) The orthodox *evasion* of the wave/particle dilemma does indeed have grave repercussions, sensed by Einstein but not by Bohr, Heisenberg and other authors of the orthodox viewpoint.

2 INADEQUACY OF EARLIER ATTEMPTS TO SOLVE THE PROBLEM

Over the years a number of attempts have been made to solve the wave/particle problem: but these are all, in one way or another, inadequate. There is Schrödinger's idea that quantum objects are quasi-classical wave-like entities, evolving deterministically in accordance his own famous quantum wave equation:

$$i\hbar\frac{\partial\psi}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\psi+V\psi.$$

But this wave interpretation of QT (WQT) suffers from the fatal flaw that it cannot do justice to the probabilistic and particle-like character of the quantum domain. There is the idea of Einstein, Popper, Landé, Ballentine and others, according to which quantum objects can be held to be quasi-classical particles obeying non-classical, statistical laws of QT. But this particle-like, statistical interpretation of QT (SQT) suffers from at least three serious defects. First, the wave-like, interference effects of QT become utterly enigmatic. There can be no explanation for wave-like effects in terms of the physical states of individual physical systems—but only in terms of statistical laws that apply to ensembles of such systems. Attempts made by Landé (1965, ch. 1) and more recently by Audi (1973, pp. 107–119) to explain the interference effects of the two-slit experiment in terms of Duane's quantum rule for objects periodic in space, seem to be untenable (Maxwell, 1975). Second, SQT is just as dependent upon classically described preparation and measurement devices as OQT is, since the statistical laws of SQT are formulated in terms of the ψ-function defined, as for OQT, in terms of preparation and measurement, and not in terms of the instantaneous physical state of quantum particles. Thus SQT is just as *ad hoc*, imprecise, restricted in scope and non-explanatory as OQT is, and for just the same reasons. Third, there is the grave objection that the positions of quantum particles of SQT just before measurement cannot be held in general to be the positions actually detected by measurement (Gardner 1972). This renders unobserved positions of quantum particles wholly metaphysical. There is the idea of de Broglie and Bohm—the idea of the 'double solution' or 'quantum potential'—according to which QT is to be interpreted as a theory about quasi-classical particles guided by a non-classical pilot wave or quantum potential, the state of which can be derived from the ψ-function of OQT. But this 'double solution' interpretation of QT (DQT) suffers from the second defect of SQT, just indicated. The state of the quantum potential in any given experimental set up is not determined by the trajectory of the individual particle in physical space, but rather by the entire experimental arrangement, the classically described preparation and measurement devices. This means DQT is just as dependent on classical physics as OQT is—and hence that it is just as *ad hoc*, imprecise, restricted in scope and non explanatory as OQT is. (For references to the interpretations of QT just discussed see Jammer 1974, Bohm and Hiley 1987.) Finally, there is Cramer's transactional interpretation of QT (TQT), according to which the quantum object, such as the photon or electron, is the outcome of the exchange of advanced and retarded waves between emitter and absorber (Cramer 1986). But insofar as quantum systems can, according to TQT, only be understood as the outcome of transactions between emitters and absorbers, TQT suffers from the same difficulty facing the other interpretations of QT just considered, namely that quantum systems can only be given quantum states with respect to other systems which are not described in terms of QT. In the case of OQT, these other systems are preparation and measurement devices; in the case of TQT they are emitters and absorbers. Perhaps much more seriously, TQT suffers from the difficulty that it requires there to be causal influences from the future to the past, as well as from the past to the future.

It is sometimes held that quantum field theory, for example quantum electrodynamics (QED) succeeds (where non-relativistic OQT fails) in solving the wave/particle problem. But this is not the case. QED, just like non-relativistic OQT, evades and does not solve the problem, in that it is a theory about the results of performing measurements, and cannot be interpreted to be a theory about the evolution of physical states of quantum fields entirely independent of preparation and measurement. All the defects which plague non-relativistic OQT, indicated above, also plague QED and other quantum field theories interpreted in the orthodox manner.

The wave/particle dilemma has been with us now for over eighty years (ever since Einstein put forward his conjecture concerning light quanta: Einstein, 1905). Some of the greatest scientific minds ever have, in one way or another, been defeated by the problem—most notably Einstein, Schrödinger, Bohr, Heisenberg, Pauli, Dirac. When an old man, Einstein remarked: 'All these fifty years of conscious brooding have brought me no nearer to the answer to the question "What are light quanta?". Nowadays every rogue thinks he knows it, but he is mistaken.' (Einstein, 1951.) In the circumstances, it is not perhaps surprising that even those concerned with interpretative problems of QT should have grown tired of the wave/particle problem—judging it no doubt to be inherently insoluble. (It is, however, let me repeat, *the* key problem we need to solve to improve our understanding of the quantum domain: the arguments of section 1 above show clearly that no version of QT can be satisfactory which fails to solve the wave/particle problem.)

Bearing all this in mind, it may seem somewhat presumptuous to claim to have *solved* the problem. This nevertheless is my claim. In this paper I not only solve the problem: I put forward a version of QT, based on this solution, which has its own consistent, distinctive quantum ontology; this version of QT is free of the defects of OQT, and in addition, leads to predictions that differ from those of OQT, as yet untested. (For earlier sketches of this version of QT see Maxwell, 1972b, 1976a, 1982, 1984, Ch. 9.) This said, I must also immediately acknowledge that many unsolved theoretical and experimental problems arise in connection with the version of QT I advocate in this paper. Far from wishing to deny their existence, quite to the contrary, the chief purpose of this paper is to highlight the existence and importance of these long neglected problems. For too long the general acceptance of a bad philosophy of science (a combination of Copenhagenism and instrumentalism) has blinded physicists to the existence of major theoretical and experimental problems of physics, having to do with the nature of quantum objects. The change of viewpoint, or of paradigm, advocated in this paper, brings these long neglected problems into sharp focus. In this sense I am advocating a new research programme rather than a new version of QT. But this, as I have argued elsewhere (Maxwell 1974, 1976b, 1984), is what the philosophy of science ought constantly to be trying to achieve: to propose new aims for research, new fruitful possibilities, neglected because of prejudice, or bad philosophy of science.

3 PROPENSITON SOLUTION TO THE WAVE/PARTICLE PROBLEM

The solution to the quantum wave/particle problem to be advocated here rests on the following two assumptions.

Assumption (I): In speaking of the *properties* of fundamental physical entities (such as mass, charge, spin) we are in effect speaking of the dynamical laws obeyed by the entities—and *vice versa*. Thus, if we *change* our ideas about the nature of dynamical laws we thereby, if we are consistent, change our ideas about the nature of the properties and entities that obey the laws.

Assumption (II): The quantum world is fundamentally probabilistic in character. That is, the dynamical laws governing the evolution and interaction of the physical objects of the quantum domain are probabilistic and not deterministic.

These two assumptions—or conjectures—are very different in character: (I) is a somewhat philosophical thesis about how we ought to conceive of the relationship between physical entities and dynamical laws quite generally, whatever the nature of the physical universe may be, while (II) is a substantial physical, or metaphysical, thesis about the nature of the quantum world.

Granted *conjectural essentialism*—a doctrine expounded and defended in the appendix below—assumption (I) follows as a triviality. (I) ought however to be accepted independently of conjectural essentialism. Almost everyone will want to concede that all physical properties are *dispositional* in character, in that they determine how things change (or resist change) in certain circumstances. (For a good, brief discussion of this point see Popper; 1959a, pp. 424–5.) Whether we have in mind rather crude common sense properties such as rigidity or inflammability, or precise, highly theoretical properties such as mass or electric charge, the result is the same: in attributing such a property to an object we thereby imply something about how the object changes, resists change, or affects change in something else, in certain circumstances. We imply something, vague or precise, about the lawful behaviour of the object. Thus in specifying the physical properties of an object we specify, vaguely or precisely, the laws that the object obeys, and *vice versa*. There is thus a one to one correspondence between the nature of (hypothetical) physical objects on the one hand, and the nature of dynamical laws on the other hand: if we change our ideas about the one we *ipso facto* change our ideas about the other—the crucial tenet of assumption (I).

As to assumption (II), this is, it may be argued, an entirely reasonable conjecture to adopt given the probabilistic character, and immense empirical success, of OQT, and given the difficulties that beset attempts to interpret OQT deterministically.

Granted (I) and (II), we are now in a position to solve the first part of the quantum wave/particle dilemma. In moving from the classical to the quantum domain there is a dramatic change in the nature of the dynamical laws taken to prevail, from *deterministic* to *probabilistic* laws (assumption (II)). This in itself demands that as we move from the classical to the quantum domain there will be a corresponding dramatic change in the kind of physical objects and properties we encounter (assumption (I)). Quantum objects and properties must differ dramatically from classical objects and properties—just because of the fundamentally probabilistic character of the quantum domain. It is thus absurd to try to understand such quantum objects as the electron and photon in terms of such inherently deterministic objects as the classical particle, the classical wave, the classical field. There is nothing inexplicable whatsoever about the fact that quantum objects such as the electron and photon differ dramatically from all deterministic classical objects (particle, wave, field). Indeed, the thing is all the other way round: granted the fundamentally probabilistic character of the quantum world, *it would be utterly inexplicable if probabilistic quantum objects did closely resemble classical objects.* Far from requiring, for comprehensibility, that quantum objects must be understandable in terms of classical notions of particle, wave or field, we must require the opposite: if the quantum world is to be comprehensible, then its objects must be understandable in terms of new probabilistic objects and properties that differ radically from classical, deterministic objects and properties.

It may be objected that classical statistical mechanics constitutes a counter example to this argument. Classical statistical mechanics is a *probabilistic* theory, and yet is about entirely classical objects—classical particles (atoms or molecules). In fact this is not a counter example. Classical statistical mechanics is not a *fundamentally* probabilistic theory: it presupposes that the basic dynamical laws are *deterministic*. Probabilism enters into classical statistical mechanics via probabilistic distributions of initial and boundary conditions in relevant ensembles of physical systems.

Einstein, Bohr, Heisenberg, Schrödinger and the other authors of QT, despite their differences, in effect agreed on one key point: if quantum objects cannot be understood in terms of the deterministic notions of the classical particle, wave or field, then this creates a severe problem for the task of developing a fully micro realistic version of QT. Bohr and Heisenberg concluded that this severe problem cannot be solved, and as a result developed a version of QT which *evades* the problem—orthodox QT. Einstein and Schrödinger were aware of the damaging consequences of this evasion, and hoped it would be possible to understand quantum objects in classical terms. What Einstein, Bohr, Heisenberg, Schrödinger et al. failed to appreciate—and what almost everyone since has failed to appreciate as well—is that the problem they all desired to solve (but which most thought insoluble) is entirely the *wrong* problem to try to solve in the first place. Failure to represent probabilistic quantum objects in terms of deterministic classical objects does *not* in itself create any kind of problem for quantum micro realism at all. Quite the contrary, a severe problem for quantum micro realism would be created if it did prove possible to represent probabilistic quantum objects in terms of deterministic classical objects. Everyone has tried to do what ought never to have been attempted in the first place. Success would have been a disaster: longstanding failure ought to be regarded as a promising sign that the quantum world may well make perfect micro realistic sense after all!

Once we appreciate what Einstein, Bohr, Heisenberg, Schrödinger et al. failed to appreciate—namely that the wave/particle problem as traditionally understood is the *wrong* problem—we can move on to formulate and solve the *right* problem. There are in effect *two* problems we need to solve in order to develop an acceptable, fully micro realistic theory of probabilistic quantum objects. First, we must specify, in general terms, the nature of entirely *unproblematic* probabilistic objects, wholly irrespective of any considerations taken from QT. Second, we must show that no difficulties lie in the way of holding that quantum objects are just such entirely unproblematic probabilistic objects (no doubt of a distinctively quantum type). We have, in short:

Problem 1: What sort of entities are unproblematic, fundamentally probabilistic objects quite generally (entirely independent of quantum mechanical considerations)?

Problem 2: Can quantum objects be construed to be varieties of a special quantum kind of such unproblematic fundamentally probabilistic objects?

Physical properties which determine how physical objects interact with one another *probabilistically* will be called here, following Popper (1957), *propensities*. Any propensity P has associated with it a number of possible outcomes $O_1 \ldots O_n$; in specifying the value of the propensity P at any instant we specify the probability p_r that outcome O_r will occur should the propensity be actualized through the occurrence of a probabilistic event at the instant in question, with $r = 1, \ldots n$, and $\sum_{r=1}^{n} p_r = 1$. (For an account of the notion of propensity that is being appealed to here, and its close analogy with deterministic, classical physical properties, see the appendix; for an account of the way this notion differs crucially from Popper's anti-essentialistic, relational notion, see Maxwell, 1976a, pp. 283–6; 1985, pp. 41–42.) Physical objects with propensities as properties will be called *propensitons*. In accordance with assumption (I), fundamentally probabilistic dynamical laws can be interpreted as specifying how values of propensities evolve, how propensitons evolve and interact.

Two kinds of fundamentally probabilistic laws need to be considered: *continuous probabilistic laws* which assert that systems evolve probabilistically continuously in time, and *discrete probabilistic laws*, which assert that systems only evolve probabilistically intermittently in time, when relevant physical conditions arise, the values of propensities (or the states of propensitons) otherwise evolving deterministically. Corresponding to these two kinds of probabilistic laws there are two kinds of propensitons, *continuous* and *discrete* propensitons.

There are, then, three kinds of dynamical theories which deserve to be regarded as equally viable from an *a priori* standpoint (other things being equal): deterministic, continuously probabilistic and discretely probabilistic theories. There is nothing intrinsically *ad hoc* or inexplicable about the instantaneous probabilistic transitions of discretely probabilistic theories: such transitions are an inherent feature of this kind of theory. Corresponding to these three kinds of equally viable theories, there are three kinds of equally viable physical entities: deterministic entities such as the classical point-particle and the classical electromagnetic field; continuous propensitons; and discrete propensitons.

The basic thesis of this paper is that electrons, photons and other quantum objects are varieties of *unproblematic discrete propensitons* (or *smearons* as they were called in an earlier paper: see Maxwell, 1982). I therefore indicate, in a little more detail, the general character of the (unproblematic) discrete propensiton.

As I have just indicated, the physical state of the discrete propensiton (and the values of the propensities it possesses) evolve *deterministically*, as long as the physical conditions for probabilistic actualization are not realized. When these latter conditions are realized, the propensiton suffers an instantaneous, probabilistic change of state, determined probabilistically by the values of relevant propensities at the instant in question. Likewise, values of propensities change instantaneously. In order to specify the nature of any (discrete) propensiton—the nature of the propensities possessed by the propensiton—three things need to be specified: (i) the deterministic dynamical laws of evolution and interaction; (ii) the precise propensiton conditions for probabilistic events to occur; (iii) probabilistic laws governing instantaneous probabilistic transitions.

One might try to visualize the evolution and interaction of the discrete propensiton in terms of the flight of a magnetized die tossed into a varying magnetic field. As the die falls the value of its propensity varies continuously and deterministically; when the die hits the table top and comes to rest, the propensity is actualized in a discontinuous, probabilistic way. This is, however, only a very inadequate model for the evolution of the discrete propensiton. The evolution of a real life, individual die can be conceived of entirely in terms of changing values of *deterministic* properties; the propensity of the die is the outcome of the statistical distribution of different initial conditions of different tosses. In the case of an evolving (discrete) propensiton, however, there is no evolution of values of *deterministic* properties—only a deterministic evolution of values of *probabilistic properties* or propensities (which is quite different). There is no *deterministic* state; only a *propensity* state. This ensures that all (discrete) propensitons are utterly unlike familiar objects, such as dice and coins, to which propensities can be attributed but which can be conceived of, more fundamentally, in terms of classical, deterministic properties.

The evolution of a genuinely (discrete) *propensiton* die would have to be conceived of in something like the following terms. The propensiton die is tossed. As the die flies through the air it is gradually transformed into six potential, virtual, ghostly dice, each with a different face uppermost, each with a different (probability) density (all equal in the case of unbiasedness), which may very well vary with time. When the six potential dice hit the table top, five vanish and one solid die remains. If the die is tossed repeatedly, the statistical outcomes are determined by the probability densities of the six virtual dice just before contact with the table top.

This, then, is the general character of the discrete propensiton. Its state evolves (i) deterministically into a smeared out range of *virtual* or *potential* states; then, when (ii) appropriate propensiton conditions arise (iii) instantaneously and probabilistically, the *virtual* states become *vacuous* except for one which becomes *actual*. Once *discrete probabilism* is conceded, this general character of the (unproblematic) propensiton is inevitable. As a special case, it is possible to envisage a kind of discrete propensiton which is such that the values of its propensities remain fixed during deterministic evolution. In this case, values of propensities will not spread out during deterministic evolution. In general, however, the propensity state of discrete propensitons will spread out during deterministic evolution—whether spatially, or in some other way. It is, in other words, not absolutely essential that discrete propensitons exhibit quantum-mechanical-type spatial smearing out, or non-locality; it is, however, entirely natural that discrete propensitons should exhibit such typically quantum mechanical features.

As a second example of a possible (unproblematic) kind of discrete propensiton, consider the following. The propensiton is in the form of a sphere, which expands at a fixed rate. The stuff of the sphere is position probability density, uniformly distributed within the sphere. The condition for probabilistic actualization to occur is for two spheres to touch. The outcome is that the two spheres collapse instantaneously into two small spheres of some minimal size, each localized probabilistically by the position probability density of each sphere. It is vital to appreciate that there is nothing inherently problematic, *ad hoc* or inexplicable about the instantaneous probabilistic collapse of the propensiton spheres (to re-emphasize a point already made). To demand that any such instantaneous, probabilistic collapse of virtual states must be *explained* in terms of some continuous evolution of state amounts to holding that only deterministic or continuously probabilistic theories are acceptable, discretely probabilistic theories being unacceptable on *a priori grounds*. Once it is conceded that these three kinds of dynamical theories are equally acceptable *a priori* (other things being equal), it is thereby conceded that the instantaneous, probabilistic collapses of propensiton states postulated by discrete probabilism are not intrinsically problematic or inexplicable—not especially in need of further explanation in terms of some continuous process.

Propensitons of this rather simple-minded type can easily be made a little more sophisticated by postulating that the position probability density is variable in space—even in a wave-like way. If the conditions for probabilistic events to occur are modified, it would even be possible to create a possible kind of propensiton which is such that an ensemble of such propensitons, passed through a two-slitted screen, creates an interference pattern of the kind created by electrons or photons.

There is nothing *ad hoc* or arbitrary about the discrete propensiton as it has just been characterized. As I show in the appendix below, as we generalize deterministic dynamical laws to become probabilistic dynamical laws, so *deterministic objects* generalize to become either *continuous* or *discrete propensitons*. The propensiton (continuous or discrete) is *the* natural generalization of the deterministic object.

So much for my solution to problem 1. My solution to problem 2 is that quantum objects can indeed be conceived of as unproblematic discrete propensitons, very roughly of the type just indicated. The two-slit experiment, for example, which so strikingly reveals both the wave-like and the particle-like aspects of electrons (or photons) can be understood in the following way. Each individual electron is in the form of a wave packet, a spatially smeared out discrete propensiton. The wave-like character of the electron propensiton is such that the absolute phase is without physical significance; only phase differences which persist through a constant change of phase of the entire wave packet are of physical significance. As a result of this, in many circumstances the wave-like character of the electron propensiton is implicit, rather than being explicit in a wave-like variation of position probability density. The propensiton states of individual electrons evolve deterministically, in accordance with the dynamical equations of QT: what evolves, however, is the propensity to interact in a probabilistic and quasi particle-like way, should the appropriate physical conditions to do so arise. The deterministic equations of QT do not of course apply to such probabilistic actualizations of propensities. When the electron wave packet encounters the two-slitted screen *either* the electron is absorbed by the screen and there is an instantaneous, probabilistic collapse of the wave packet, *or* the electron wave packet passes through both slits. Granted the latter then, on the other side of the screen, the implicit wave-like character of the propensiton state of the electron leads to interference (as a result of phase differences which cannot be eliminated by any constant global change of phase): the wave-like character of the propensity state of the electron becomes explicit in an interference-like variation of position probability density. The wave packet then encounters the photographic plate and interacts with all available silver bromide molecules. The physical condition for a propensiton or wave packet probabilistic collapse are then realized: abruptly, the electron continues to interact in a highly localized way with just one silver bromide molecule (or crystal) in such a way as to create a developable dot of silver on the photographic plate. The position of the dot is probabilistically determined by the interference pattern of position probability density of the electron propensiton just before the wave packet collapse. As a result, in the case of an ensemble of similarly prepared electrons with the same momenta, and therefore the same wavelengths, the developable dots on the photographic plate fall into the characteristic observed interference pattern—mirroring the interference pattern in position probability density of each individual electron propensiton just before probabilistic localization occurs.

It deserves to be noted that the electron, conceived of as a distinctively quantum mechanical kind of discrete propensiton, exhibits particle-like features in two ways. First, a particle-like aspect is exhibited whenever, as a result of a probabilistic propensiton collapse, the electron is detected in a localized way as a dot on a photographic plate, or as a trail of ionized molecules or water droplets in a Wilson cloud chamber. Second, a particle-like aspect is exhibited in the dynamical character and behaviour of the electron propensiton. The field of force created by the electron propensiton corresponds, not to a classical charged point-particle, but rather to a *superposition* of charged point-particle states. Consider Schrödinger's time-dependent equation for two particles:

$$i\hbar\frac{\partial}{\partial t}\psi(r_1r_2t)=-\frac{\hbar^2}{2m_1}\nabla_1{}^2\psi(r_1r_2t)-\frac{\hbar^2}{2m_2}\nabla_2{}^2\psi(r_1r_2t)+V(r_1r_2)\psi\ (r_1r_2t).$$

In the case of two non-relativistic electrons interacting by means of their electrostatic force alone, the potential function $V(r_1r_2)$ becomes e^2/r_{12}, and the last term of Schrödinger's equation, $e^2/r_{12}\psi(r_1r_2)$, can be regarded as specifying a *superposition* of electrostatic forces between all pairs of coordinates r_1, r_2, for which $|\psi(r_1r_2t)|^2dr_1dr_2>0$.

There are now a number of tasks I need to accomplish to transform this solution to the wave/particle dilemma into a fully fledged propensiton version of QT (PQT). First, I need to specify how OQT is to be modified so that it becomes PQT—a version of QT which is exclusively about quantum propensitons evolving and interacting in space and time, in the first instance entirely independently of preparation, measurement and classical physics (PQT thus being free of the seven defects which plague OQT). Second, and most important of all, I must specify the precise, necessary and sufficient, quantum mechanical conditions for probabilistic events to occur—for quantum propensitons to suffer instantaneous collapse. Third, I need to specify precise probabilistic laws governing quantum propensiton collapse. Fourth, I must show that PQT recaptures all the empirical success of OQT, even though PQT eschews all reference to observables, measurement and classical physics. Fifth, I need to indicate crucial experiments capable of deciding between OQT and PQT. These five points are taken up in turn in the remaining sections of the paper.

4 FROM ORTHODOX TO PROPENSITON QUANTUM THEORY

OQT consists of two parts. On the one hand there are the dynamical equations, such as Schrödinger's time-dependent and time-independent equations, the Klein-Gordan and Dirac equations, and the equations of quantum field theory. On the other hand there are the interpretative postulates of OQT, which interpret the ψ-function in terms of measurement. These generalize Born's 1926 postulate (Born, 1926, 1927), and may be taken to assert:

(1) If a measurement of observable A is performed on a system (or ensemble of systems) in a state ψ, then the probability of obtaining a value between a_r and $a_{r+dr}=|(\alpha_r,\psi)|^2dr$, where (a_r) and (α_r) are eigenvalues and eigenvectors of the Hermitian operator $\hat{A}$ corresponding to the observable A.

PQT retains the dynamical equations of OQT but rejects the interpretative postulate (1). Instead of interpreting ψ as containing information about values of *observables*, about the outcome of performing *measurements* on the system (or ensemble of systems) in question, PQT rather interprets ψ as specifying the actual physical state of the individual quantum system in physical space and time, even in the absence of preparation and measurement. All quantum systems are conceived to be discrete propensitons (as indicated in the last section). ψ is interpreted to contain information about the values of various quantum propensities of quantum propensitons, such as position, momentum and energy probability density, and angular momentum or spin states. In specifying how ψ evolves in time, the dynamical equations of QT specify how values of these propensities evolve deterministically, just as long as no probabilistic events occur. From the outset, and necessarily, the *scope* of the dynamical equations of QT is restricted to the deterministic evolution of quantum propensitons. Whenever quantum propensities, such as position, momentum or energy density, are probabilistically actualized then, at that instant, deterministic dynamical equations do not apply. (It is this restriction of the scope of the dynamical equations of QT, basic to the whole propensiton idea, which ensures that any precisely formulated version of PQT must differ experimentally from OQT, at least in principle.) Instead of the generalized Born postulate (1) of OQT we have, within PQT, postulates which specify the precise quantum propensiton conditions for probabilistic events to occur, and the precise instantaneous and probabilistic changes of propensiton state that result. All quantum measurements will turn out to be no more than special cases of a kind of probabilistic process occurring naturally, throughout the universe. PQT enables us to derive Born's postulate from purely quantum mechanical postulates, without any assumption being made concerning observables, measurement or classical physics. Stable macro objects and macro phenomena, obeying approximately classical laws, emerge naturally, according to PQT, as the outcome of vast numbers of quantum propensitons interacting with one another in a probabilistic manner. (Earlier sketches of PQT are to be found in Maxwell 1972b, 1976a, 1982, 1984 Ch. 9, 1985.)

It might seem that the *complex* character of ψ constitutes a serious obstacle to interpreting it as specifying *real* values of propensities of real propensitons. But this is not the case. We may take $|\psi|^2dV$ to specify the real value of the propensity, position probability density, within each dV. Analogous remarks hold for momentum and energy probability density, and spin. In this way ψ, a complex function of space and time, is interpreted to attribute real values of quantum propensities to quantum objects in physical space and time. (This does not reintroduce the notion of measurement. Quantum propensities presuppose probabilistic *localizations*, but not *measurements*: see Maxwell, 1976a, pp. 661–3.) It would seem that ψ is complex in order to do justice to the often implicit wave-like character of quantum systems, alluded to above, in the last but one paragraph of section 2. Thus a quantum system in an eigenstate of momentum has a definite wavelength associated with it, even though position probability density is constant in space. If ψ is complex, this state of affairs is easy to depict. The wave character of the quantum state may be represented by: $e^{i\alpha} = \cos\alpha + i\sin\alpha$; since $|e^{i\alpha}|^2 = 1$, the square of the amplitude can be constant and the wave character of the state still exist. The time-dependent Schrödinger equation specifies how such an implicit wave-like feature of a quantum state can, in certain circumstances, become explicit as an interference-like variation of position probability density in space—as a result of diffraction, for example.

It might seem that the fact that, for 2 (or n) interacting systems we need to resort to a ψ-function in six (or 3n) dimensional configuration space delivers a fatal blow to the propensiton interpretation of QT. How can such a ψ-function be interpreted as specifying the real physical states of 2 (or n) objects in 3 dimensional physical space? In order to carry through such an interpretation, we must first appreciate that n interacting quantum objects do not have independently specifiable propensity states: only the composite object as a whole has a definite propensity state. The propensities of this composite propensiton, in 3 dimensional physical space, need to be understood as follows. Consider position probability density. For an n particle system in a state ψ, this is represented by $|\psi|^2 dr_1 \ldots dr_n$, and is to be understood as determining: the probability of particle 1 being available for a probabilistic interaction in dr_1, particle 2 in $dr_2 \ldots$ and particle n in dr_n (for all possible values of $dr_1 \ldots dr_n$). Instead of interpreting $|\psi|^2 dr_1 \ldots dr_n$ as assigning a probability to a small region in 3n dimensional configuration space, we interpret it as assigning a probability to n small regions $dr_1, \ldots dr_n$ in 3 dimensional physical space. The *value* of this propensity cannot be uniquely specified for particle 1 in region dr_1 independently of the other particles: as $dr_2, \ldots dr_n$ are moved through space, the overall probability of particle 1 being available for interaction in the *fixed* region dr_1 (and the other particles being available in $dr_2, \ldots dr_n$) will *vary* as well. What exists potentially in one small spatial region at an instant depends, in this way, on what exists, potentially, elsewhere—a feature of the quantum world not encountered within classical physics. To say this, however, is just to say that the n interacting particles do not have distinct quantum states, but only have a joint, quantum-entangled state as a whole. In order to specify how the value of the n-fold position probability density of the n-particle system, and the values of other such propensities, evolve in physical space and time, it is convenient to resort the mathematical fiction of a ψ-function with a unique value at each point in 3n dimensional configuration space. This is to be interpreted physically, however, as assigning a unique value to any n points in 3 dimensional physical space.

This physical interpretation of the propensities of interacting and composite quantum objects may well be of special significance when it comes to the question of how the conditions for probabilistic events to occur are to be specified—as we shall see below.

5 VIRTUAL, POTENTIAL AND ACTUAL PARTICLES

According to the version of PQT to be considered here, probabilistic events occur (with the actualization of quantum propensities) when and only when, as a result of inelastic collisions or decay processes, new *actual* (as opposed to merely virtual) particles are created—all quantum measurements that detect systems being merely special cases of particle creation.

One reason why this postulate seems worthy of serious consideration can be put like this. As long as quantum objects interact with macroscopic objects in an *elastic* fashion as when electrons are diffracted through a crystal or two-slitted screen, no probabilistic localization seems to occur. It is when quantum objects interact *inelastically* in a highly *localized* fashion, to create new particles or ionized molecules, that probabilistic wave packet collapse seems to occur. All quantum measurements that actually *detect* quantum systems (and do not merely *prepare* quantum states) must involve some such *inelastic*, particle-creating process—usually millions of such processes—simply to produce a permanent record (necessary for measurement to have taken place). Granted that *detection* is (in general) a sufficient condition for a probabilistic event to occur, and granted we seek some elemental quantum condition for the occurrence of probabilistic events, it seems not unreasonable to conjecture that creation and annihilation of particles—whether elementary or composite—is the proper necessary and sufficient quantum condition for the occurrence of probabilistic events (quantum measurements thus exemplifying physical processes that occur in Nature all the time).

One immediate objection which may be made to the above proposal is that it requires something which does not exist—an absolute distinction between virtual and actual particles. But to this we can reply that the distinction can be drawn quite straightforwardly as follows. Virtual particles are particles whose persistence is constrained by a combination of uncertainty relations and conservation principles, whereas actual particles are subject to no such constraint.

Strictly speaking, within the framework of orthodox quantum field theory, it is possible to distinguish *three* kinds of status for particles: *virtual, potential* and *actual* (an important point that does not seem to have been made explicitly hitherto in the literature). Whereas virtual particles are such that their persistence is restricted by conservation principles, both potential and actual particles are, according to OQT, subject to no such restriction. *Potential* particles arise whenever, as a result of inelastic collisions, two or more interaction channels result, different kinds or numbers of elementary particles being associated with each channel. The channels are *alternative possibilities* rather than *actualities* (in that measurement can detect only *one* channel). *Actual* particles, on the other hand, are particles which are not a part of any such superposition of alternative possible particle states.

These distinctions can be clarified by means of the following example. Consider a photon sufficiently energetic to create an electron/positron pair. In the vacuum such a photon may be regarded as creating virtual electron/positron pairs: conservation of momentum and energy ensures, however, that such pairs do not persist. Here, the electron/positron pairs are irredeemably *virtual*, whereas the photon is *actual*. If however the photon encounters a nucleus, a persisting electron/positron pair becomes possible, since the nucleus can carry off energy and momentum in such a way that both energy and momentum are conserved. Suppose there is a probability $= 1/2$ that the photon will subsequently be detected, probability $= 1/2$ that the electron/positron pair will be detected. OQT predicts that the system persists as a superposition of these two alternative possible channel states until a measurement is made. Here, the electron/positron pair on the one hand, and the photon on the other hand, are *potential* rather than *actual* particles. A measurement which establishes the non-existence of the photon converts the electron/positron pair from *potential* to *actual* status. Likewise a measurement which detects the electron converts the status of the positron from *potential* to *actual*.

It is of vital importance to appreciate, here, that a system which consists of an *actual* photon and an *actual* electron/positron pair is quite different from a system consisting of a *potential* photon and *potential* electron/positron pair. In the former case appropriate measurement detects *both* photon *and* electron/positron pair: in the latter case measurement can only detect *either* the photon *or* the electron/positron pair. In the former case there is sufficient energy and momentum for *both* photon *and* electron/positron pair to exist: in the latter case there is only sufficient energy and momentum for *either* the photon *or* the electron/positron pair to exist.

According to OQT, superpositions of different *potential* particle states can persist without limit. Thus, in the above example, the superposition of the potential photon and potential electron/positron pair persists indefinitely, as long as no measurement is performed. According to OQT, indeed, the *only* way *potential* particles can become *actual* is as a result of measurement. Even more paradoxically, according to OQT, the number of potential particles increases without limit, in the absence of measurement. The number increases every time there is an inelastic collision with more than one interaction channel outcome. This is the case even if we allow that particles can be annihilated as a result of interactions or decay—since such processes do not decrease the number of alternative potential particles. Strictly speaking, of course, particles cannot be annihilated in this way, since there must always be a non-zero amplitude for the particles not to be annihilated—which only makes the situation worse.

We have here a new difficulty facing OQT, a new 'potential particle' paradox, somewhat analogous to Schrödinger's cat paradox, but with this

PG152

difference: whereas Schrödinger's paradox arises when OQT is applied to the macro domain (e.g. to a cat), the potential particle paradox arises entirely within the quantum micro domain itself.

PQT solves this new potential particle paradox at a stroke: PQT demands that potential particles persist only for as long, roughly, as corresponding virtual particles do, and then either cease to exist or become actual, entirely independent of measurement.

6 PARTICLE CREATION AND ANNIHILATION AS THE QUANTUM CONDITION FOR PROBABILISTIC EVENTS TO OCCUR

Suppose that, as a result of an inelastic interaction between wave packets, alternative possible interaction channels are created, each channel having its own distinctive clutch of potential particles. In the case of the $\gamma \rightarrow e^+ + e^-$ interaction discussed above, there are *two* channels, the first consisting of photon and nucleus, the second consisting of e^+/e^- pair and nucleus. Whereas OQT asserts that such channels persist indefinitely until measurement detects one or other channel outcome, PQT asserts:

Postulate (2): A *sufficient* condition for the superposition of channel states to have decayed probabilistically into one or other channel state, with its distinctive clutch of potential particles becoming actual, is that the interaction responsible for creating the potential particles of the different channels ceases, due to the spatial separation of interacting wave packets.

Thus, according to (2), for the $\gamma \rightarrow e^+ + e^-$ case, the superposition of potential photon state and potential $e^+ + e^-$ state is *either* the actual photon state (with probability $= \frac{1}{2}$) *or* the actual $e^+ + e^-$ state (with probability $= \frac{1}{2}$) once the initial photon has separated spatially sufficiently from the nucleus.

The basic idea of (2), and of all more precise postulates to be formulated below, is that interactions proceed entirely in accordance with the dynamical equations of OQT with this one exception: whereas OQT asserts that different channel outcomes of inelastic interactions persist as superpositions, PQT asserts that such superpositions decay spontaneously and probabilistically into one or other channel state. We have here, then, a possible propensiton alternative to orthodox quantum field theory (OQFT), namely propensiton quantum field theory (PQFT)—or relativistic quantum propensiton theory, as it ought perhaps to be called. The difference between OQFT and PQFT can be illustrated by means of (a perhaps somewhat illegitimate use of) Feyman diagrams of the $\gamma \rightarrow e^+ + e^-$ process already discussed (see diagram 1).

In the space-time region indicated by the dotted circles (the region of the interaction $\gamma \rightarrow e^+ + e^-$), OQFT and PQFT agree: what exists is a superposition of the four processes indicated, and processes represented by all higher order diagrams (here neglected). In space-time regions outside, and after, the region

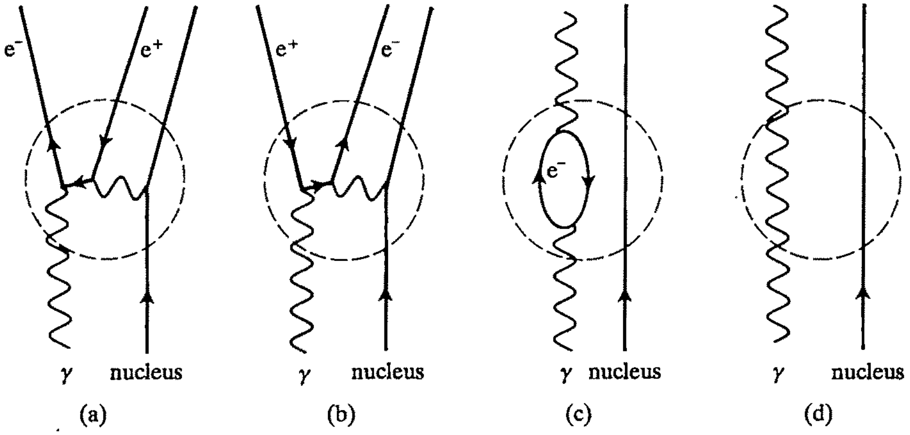


Diagram 1

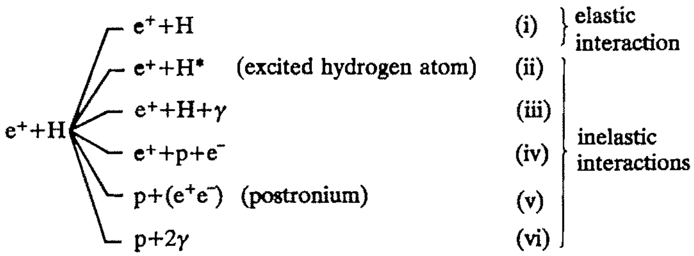

e⁺+H
e⁺+H
e⁺+H* (excited hydrogen atom)
e⁺+H+γ
e⁺+p+e⁻
p+(e⁺e⁻) (postronium)
p+2γ
(i)
(ii)
(iii)
(iv)
(v)
(vi)
elastic interaction
inelastic interactions

Diagram 2

indicated by the dotted circles, OQFT and PQFT disagree. According to OQFT, the superposition of (a), (b), (c) and (d) persists; according to PQFT, the individual system is *either* in the state that is the outcome of (a) and (b), *or* in the state that is the outcome of (c) and (d). If we neglect this difference, then OQFT and PQFT agree. In particular, the two theories agree concerning the evolution of systems composed of persistently interacting virtual or potential particles.

Postulate (2) has its limitations. It does not specify precisely when, where and how the decay of channel states occurs. It is restricted to spatially finite wave packets, and does not seem to be applicable to decay processes. It is put forward as a first approximation to a more satisfactory postulate, to get the discussion of possibilities underway. In the next section, a number of rival postulates will be considered, each specifying more precisely, and in a more generally applicable way, how, when and where probabilistic collapse of channel superpositions occurs. The rest of this section is devoted to a discussion of the crucial question: How are interaction channels to be distinguished?

The point is this. The basic idea behind (2) is open to two interpretations, namely:

(2A): Probabilistic events occur when and only when *elementary* particles are created or annihilated.

(2B): Probabilistic events occur when and only when new particles are created *whether elementary or composite*, as long as the total rest mass of particles in different potential outcomes is different.

(2A) and (2B) are different versions of postulate (2), in that they interpret the key notion of 'interaction channel' differently, as this notion figures in (2). Thus, according to (2A), there are *two different* interaction channels (emerging from the same interaction) if and only if there are different potential elementary particles in each channel (including different numbers of the same kind of elementary particle). According to (2B), on the other hand, there are *two different* channels if and only if there are different potential particles in each channel, whether *elementary or composite*, the total rest mass of particles in different channels being different.

The difference between (2A) and (2B) can be clarified by considering the interaction in diagram 2.

There are here, according to (2A), just *three* alternative channel outcomes (only one of which survives probabilistic collapse) namely [(i), (ii), (iv), (v)], [(iii)], and [(vi)]. According to (2B), on the other hand, (i), (ii), (iv) and (v) are *different* channel outcomes, since each consists of different composite particles with different total rest masses. (The mass of the hydrogen atom H is slightly less than the sum of the masses of its constituents p and e^-.) Whereas (2A) predicts that the superposition of states (i), (ii), (iv), and (v) persists indefinitely, (2B) predicts that this superposition decays spontaneously and probabilistically into one or other state. However, (i) and (iii), which are distinct potential particle states for (2A), are the same for (2B), since they do not differ in total rest mass.

Postulates (2A) and (2B) give rise to different problems, and will be discussed in turn.

Postulate (2A), on the face of it by far the more attractive possibility, is confronted by at least four problems. First, in order to be acceptable, (2A) must predict that probabilistic events occur for all types of quantum measurements that actually detect particles. It would seem, however, that measurements can be performed without the creation of new elementary particles—as when electrons are detected by means of the ionization of molecules in a cloud or bubble chamber, or by means of chemical processes associated with photography. Second, there are straightforward coherent quantum states of the electromagnetic field to which no definite photon number can be assigned: (2A) however seems to require that photon number is not ambiguous in this way. Third, there is a problem concerning the identity of photons. When a photon interacts with an electron, as in Compton scattering, or with an atom, so that it is absorbed and emitted by the atom, under what circumstances does such a process constitute (a) an elastic interaction of one photon (there thus being no probabilistic event) or (b) an inelastic interaction involving the annihilation of one photon and the creation of a second photon (there being in this case a probabilistic event)? Fourth, what is the status of elementary particles in hadrons, nuclei and other persisting composite quantum objects? In such cases it is difficult to see what alternative we have to holding that the composite object is, internally, a superposition of virtual or potential particle states, the *composite* object, nucleon, nucleus or whatever, being the *actual* particle.

The first problem is not too serious. According to (2A), whenever a measurement of the kind indicated is performed, an atom or molecule (at the very least) goes into a superposition of two states (as when a molecule is in a superposition of the non-ionized and ionized state, or silver and bromide atoms are in superpositions of the atomic and molecular states). These states interact with photons in quite different ways, creating and annihilating photons quite differently. According to (2A), it is the creation and annihilation of such secondary photons which leads the superpositions of atomic and molecular states to collapse probabilistically into one or other state—thus creating the definite observed outcomes of measurement. The second objection is much more serious, and may indeed suffice to demolish (2A). In order to rescue (2A) from this objection it would seem to be necessary to restrict (2A) to contexts in which photon number is unambiguously defined (unless a way can be found to reformulate (2A) so that precise conditions for probabilistic events to occur are specified even when the number of photons being created or annihilated is, within certain limits, ambiguous). The third objection can be overcome by insisting that, for a probabilistic event to occur, photon annihilation must be *actual* and not just *virtual* or *potential*. Granted that the energy of the photon is dE, actual annihilation only occurs if the photon is annihilated for a period $dt >> \hbar/dE$. The fourth objection can be met with the reply that even if quarks in hadrons and nuclei have a virtual or potential status, the composite particle alone being *actual*, nevertheless it can still be the case that probabilistic events occur when created potential elementary particles emerge from interaction cells (or tubes) in the way indicated above.

I turn now to a consideration of postulate (2B). This postulate avoids entirely the four problems that beset (2A), but at the price of creating two new difficulties. How can the *bound* state of N elementary particles be distinguished from *excited* bound states, or indeed from unbound states, in a sufficiently precise, general, non-arbitrary way to provide an adequate basis for specifying precise conditions for probabilistic events to occur, as required by (2B)? What rationale can there be for holding that the creation or annihilation of a bound state leads to a probabilistic event when all that is involved is a rearrangement of elementary particles?

In order for N particles to constitute *one* composite particle throughout some space-time region, in the sense required by (2B), the following must be satisfied. First, it must be possible to factorize the state ψ_N of the particles at any instant into two parts: on the one hand there is the 'internal' state ψ_{int}, formulated in terms of the relative positions of the N particles, and corresponding to some *definite energy level*; and on the other hand there is the 'external' state ψ_{ext}, a wave function of the centre of mass of the system in three dimensional space, and varying with time. Second, this state must persist for long enough for the fixed internal energy level to be distinguishable from other possible energy levels. The bound particle then has a mass equal to the mass of the N particles minus the mass corresponding to the binding energy associated with the energy level in question. This state is distinct from bound states that are superpositions of energy levels, and distinct from interacting, unbound states. (2B) asserts that a superposition of (a) any such composite particle state (with a definite internal energy level) and (b) other states decays probabilistically into *either* (a) *or* (b).

A preliminary rationale for adopting (2B) rather than (2A) quite independent of the problems which confront (2A) can be given as follows. In appropriate circumstances, the laws of QT appear to be just as simply and straightforwardly applicable to complex phenomena as to elementary phenomena (in striking contrast to dynamical theories, whose applications usually become rapidly much more complicated as phenomena cease to be elementary). This feature of the laws of QT is illustrated by the elementary quantum mechanical behaviour of complex, composite particles, in appropriate physical circumstances. It is illustrated by the phenomenon of phonons—elementary quanta with a highly complex molecular composition. And it is illustrated by such macroscopic quantum phenomena as superfluidity. QT appears to be a general mechanics concerned with the way energy, momentum, frequency and wavelength are inter-related (at levels comparable to Planck's constant) however elementary or complex the objects may be which are undergoing motion and vibration. Once we adopt this standpoint, it becomes clear that (2B) is more appropriate than (2A)—as long as (2B) can be formulated with absolute precision and generality.

The intended basic idea behind both (2A) and (2B) is the entirely natural and plausible postulate that:

(3): Probabilistic events occur when and only when quantum objects of some characteristic kind are created or annihilated.

Postulate (2A) clearly captures the basic idea of (3): but does (2B)? How can the creation of new bound states of the *same* elementary particles be held to constitute the creation of new kinds of quantum objects?

That (2B) does capture the basic idea of (3) can be argued for as follows. Consider a system of N elementary particles. This system in the form of a bound state, with a stationary internal state, is indeed a different kind of physical object in physical space and time from the object made up of the same elementary particles interacting in a non-stationary way, or not interacting at all. In the case of the stationary bound state, the N elementary particles have an external quantum state in physical space *as a whole*. On the one hand there is the external state of the composite particle; this can be conceived of as a propensiton state of a physical object in physical space, analogously to the way in which the propensiton quantum state of an elementary particle can be conceived: on the other hand there is the stationary internal state of the composite particle. The bound state behaves as if it were an elementary particle, without internal structure, with overall spin and charge properties that may be quite different from those of its constituents, as long as it interacts with its environment sufficiently non-energetically for its internal structure to remain unaffected. The case of the N elementary particles interacting in a non-stationary way is different. Here the quantum state of the entire system can, it is true, be represented as the product of a state vector ψ_{ext} that is a function of the coordinates of the centre of mass of the system in physical space, and a state vector ψ_{int}, that is a function of the relative coordinates of the N elementary particles. But in the first place this depends on Galilean invariance: such a factorization is not in general possible within relativistic QT. Secondly, ψ_{ext}, which may in a sense be regarded as representing the external quantum state of the system as a whole in physical space, does not represent the propensity state of an actual physical object: it represents only the centre of mass of the N elementary objects. This can in general only be detected *via* the detection of the

N distinct particles. The composite particle, on the other hand, can be detected as a single particle; it may interact with other systems as a whole physical object in its own right, as long as its internal structure remains unaffected. The composite particle with a stationary internal state is, in short, a different kind of object from the system of N interacting elementary particles without a stationary internal state. (That these two cases represent fundamentally different kinds of objects in physical space is strikingly apparent within the framework of PQT, according to which all quantum state vectors must be interpreted as specifying propensity states of objects in three dimensional physical space. When this restriction is relaxed, as it is within the framework of OQT, and the state vector of N particles can be interpreted in terms of a wave function in 3N configuration space, the fundamental character of the above distinction may well be overlooked.)

There is, it must be emphasized, nothing unorthodox about the thesis that composite particles have quantum states *as a whole*. The fact that the two-slit experiment or the Stern-Gerlach experiment can be performed with *atoms* demonstrates convincingly that composite quantum objects can have wave packets in physical space *as a whole*, as if the composite object were elemental, without internal structure. In the case of the proton and neutron, indeed, it was not realized for some decades that these objects have a complex inner structure. As Gottfried and Weisskopf (1986, p. 14) remark: 'A system in its ground state can be considered to be endowed with fixed, unchanging properties as long as the energy exchanges with its environment are much less than the difference ΔE between the first excited state and the ground state. It then acts like an 'elementary' particle with fixed properties. Among those properties we mention its spatial extension and symmetry, its angular momentum, and its magnetic and/or electric multipole moments. The system changes these properties only if it is excited to higher states; whenever it returns to the ground state the system regains the properties typical of that state.'

Where PQT (incorporating (2B)) differs from OQT is not in its assertion that composite objects have external quantum states *as a whole*, but rather in its assertion that non-interacting superpositions of such states and non-bound, non-stationary states of the same elementary particles *decay spontaneously into one or other state.*

I give now an argument in support of (2B). Two closely related principles of QT are:

(4): A system in a superposition of energy eigenstates differing by ΔE oscillates with frequency $= \Delta E/\hbar$.

(4*): A system in a superposition of different potential particle states with internal energies differing by ΔE oscillates with frequency $= \Delta E/\hbar$.

Oscillations of the nitrogen atom of the amonia molecule illustrate (4): oscillations of the K_S and K_L (potential) particle states associated with neutral kaon decay illustrates (4*). In the physics literature, these two kinds of cases are treated as if they exemplify one and the same quantum mechanical principle: see Feynman et al. (1965, Chs. 8–11); Frauenfelder and Henley (1974, pp. 214–221). In fact, even though closely related, (4) and (4*) are not quite the same. Whereas (4) requires that the system be in a superposition of distinct eigenstates of energy, (4*) does not. (4*) asserts that a superposition of different potential particle states, with rest masses or internal energies differing by ΔE, oscillates between these particle states even if the overall system is (nearly) in an eigenstate of energy.

Granted that (4*) is an universally valid principle, corroborated by neutral kaon decay and other phenomena, we now have the following remarkable result. Whenever inelastic collisions create two or more different potential particle states then, according to OQT, oscillations must persist, in accordance with (4*). On the other hand, as wave packets associated with these different particle states separate spatially, oscillations become physically impossible in that probability density cannot oscillate between the states. *The superposition of alternative potential particle states has become an impossible state.* OQT, implying both (4*) and the persistence of superpositions of potential particle states, leads to a contradiction. Granted that (4*) is upheld, the orthodox thesis that superpositions of different potential particle states persist must be rejected. (4*), in brief, implies (2B). This constitutes a strong argument in support of (2B). In what follows, (2B) rather than (2A) is presupposed.

7 POSSIBLE LAWS GOVERNING PROBABILISTIC EVENTS

In order to give precision to PQT (and to the nature of quantum propensitons) two laws governing probabilistic events need to be formulated. The first kind of law, like the one formulated in the last section (whether as postulate (2A) or (2B)) specifies precise necessary and sufficient quantum conditions for an instantaneous probabilistic event to occur, in terms of the state ψ of the system. The second kind of law specifies how, given that a probabilistic event occurs, the state of the system ψ determines n possible outcomes ψ_r ($r = 1 \ldots n$) and assigns probabilities p_r to each ψ_r (with $\sum_{r=1}^{n} p_r = 1$).

It is assumed here that the first kind of law is of the following general form: whenever a certain kind of superposition of states $\psi = \sum_{r=1}^{n} C_r\psi_r$ comes into existence, then this superposition decays probabilistically into one or other ψ_r. Granted this, then the dynamical laws of QT, together with this first kind of law, determine the second kind of law. The form of the second law will be:

(5) Probability $[\psi(t) \rightarrow \psi_r(t)] = |C_r|^2$, where $\psi(t) = \sum_{r=1}^{n} C_r\psi_{r(t)}$, $r = 1...n$, and

$$\sum_{r=1}^{n} |C_r|^2 = 1.$$

It suffices, then, to consider only candidates for the first kind of law. I now consider six possibilities.

Suppose that, as a result of inelastic wave packet collision, or decay, new potential particles are created in a region of what may be called the configuration space-time of the total system—the outcome being a number of different interaction channels. This creation interaction region of configuration space-time—the region within which and throughout which creation of potential particles can occur—can be interpreted as specifying a creation region in physical space-time. The creation region in configuration space-time can be subdivided into N creation cells, each of which specifies a creation cell in physical space-time, $dr^3 \cdot dt$. For simplicity, we suppose that there are just two possible channel outcomes—for example the photon, and the e^+/e^- pair, in the $\gamma \rightarrow e^+ + e^-$ interaction discussed above. To a first approximation, the dimensions of the creation cells are given by $dt = \hbar/dE$, $dr = cdt$, where $dE = dMc^2$, and dM is the difference in total rest mass of the potential particle states. (In the case of the $\gamma \rightarrow e^+ + e^-$ interaction, dM = rest mass of the e^+/e^- pair.) We then have:

(6A): Within each cell there exists a superposition of the elastic channel state and the inelastic channel state created within the cell (e.g. a superposition of the potential photon and potential e^+/e^- pair states). Outside the cell, this superposition decays spontaneously and probabilistically into *either* the elastic channel state (the photon) *or* the inelastic channel state (the e^+/e^- pair). This is the *only* way in which the evolution of quantum states differs from that which is specified by OQT (or by OQFT). If all *inelastic* channel outcomes, created in all cells, fail to become actual, no probabilistic localization takes place, and the elastic channel outcome is a superposition of outcomes of interactions of all possible interaction cells—whether *creation* cells or not (with possible contributions from *virtual* inelastic channel states). If however one inelastic channel outcome, created within one cell, becomes actual, then there is at that instant a probabilistic change of state: all other channel outcomes, associated with all other interaction cells, are instantaneously annihilated, and the entire system emerges in a highly localized way from the one cell as the inelastic channel outcome. Thus, in terms of the $\gamma \rightarrow e^+ + e^-$ interaction, if one potential e^+/e^- pair, created within one cell, becomes actual, then there is at that instant a probabilistic change of state: the photon, and all other potential e^+/e^- pairs are instantaneously annihilated, and the actual e^+/e^- pair, and the nucleus, emerge in a highly localized way from the specific cell in question.

If the e^+/e^- pair is equally likely to be created in any one of N cells, the probability it is created in any specific cell $= 1/(2N)$, given probability $\gamma \rightarrow e^+ + e^-$ as a whole $= 1/2$.

It may be that for certain interactions, such as neutral kaon decay for example, creation interaction cells have more the form of time-like tubes rather than cells, the position probability density of the persistently interacting potential particles within each tube becoming progressively attenuated as, roughly speaking, potential particles pass through the surface of the tube with the passage of time. In this case, as created potential particles pass through the surface of the interaction tube, *either* they are annihilated, *or* they become actual, instantaneously and probabilistically, in a localized fashion. In short, the case of the interaction tube does not introduce anything essentially new that is not already found in the case of the interaction cell.

In the case of interactions that lead to more than two different channel outcomes (as in the $e^+ + H$ interaction discussed above), the spatial part of each creation cell needs to be conceived of as a sort of interaction onion, different channel outcomes becoming progressively annihilated as each spherical shell of the onion is reached, the energy difference between channels becoming progressively less and less.

Whether quantum wave packet collision theory can be modified so as to incorporate postulate (6A) is, for me at least, an open question. One elementary consideration which suggests that it may not be possible to modify quantum wave packet collision theory so as to incorporate (6A)—and so as to be compatible with experimental results—is the following.

As a result of insisting that actual particle creation takes place, to a first approximation, in some specific small region of space of size $dr = c\hbar/dE$, (6A) places a severe restriction on the form of subsequent quantum states of the created particles. It is possible that inelastic channel states, initiated in this spatially highly restricted way, cannot yield the actual results of scattering experiments successfully predicted by OQT. It must be remembered that OQT rejects (6A). Given two spatially extended wave packets colliding in such a way as to create new particles then, according to OQT, in the absence of measurement, the new particles are *not* created in some specific small region of space (but rather in a superposition of such small regions). This difference between OQT, and PQT based on (6A), may ensure that scattering experiments already performed confirm OQT but *refute* this version of PQT.

If so, (6A) can, it seems, be modified to avoid such a refutation. Let us suppose that for any given type of interaction there is a 'minimal interaction region' dR, which is the smallest spatial region to which wave packets of created actual particles can be restricted, in order that PQT yields the predictions of OQT for experiments performed on any one channel. We then have:

(6B): Wave packet collisions, or decays, which create new particles, occur as

(6A) asserts except that the outcomes of neighbouring cells persist and interfere with one another throughout minimal interaction regions, dR: for any particular dR, *either* inelastic channel outcomes emerging from dR spontaneously vanish *or* one inelastic channel becomes *actual* and all other states of the system vanish.

If dR = dr, (6B) becomes (6A). If dR = R, the entire region of interaction of the wave packets, then (6B) becomes:

(6C): The necessary and sufficient condition for superpositions of channel states to decay into one or other channel state is that the interaction responsible for creating the different channels ceases, due to spatial separation of interacting wave packets.

(6C) is simply postulate (2) of the last section reformulated as a *necessary* condition—in addition to a sufficient condition—for probabilistic actualization to occur.

Some defects of (6C) or (2), already indicated, can be overcome by reformulating the postulate as follows. Given that two long wave packets interact in a potentially inelastic way, at any given instant, the state ψ of the total system occupies three spatial regions. There is R_1, the region which contains that part of ψ which has not yet interacted (sufficiently strongly to create the inelastic outcomes); there is R_2, the region of interaction; and there is R_3, the region which contains that part of ψ that has ceased to interact. (6D) asserts that as the interaction proceeds, at some instant the total state ψ decays probabilistically into one or other of the outcome states occupying R_3. The probability P_t that this has occurred at some instant up to time t is given by $\int_{R_3}|\psi(t)|^2dV$. We have:

(6D): $P_t=\int_{R_3}|\psi(t)|^2dV$.

The fifth possibility to be considered is:

(6E): The necessary and sufficient condition for the superposition of different channel outcomes of an interaction confined to some region R_1 to collapse probabilistically into one or other channel state is that the interaction products begin to interact inelastically with some new system in some region R_2, there being no spatial overlap between R_1 and R_2.

There is one more postulate that must be formulated because of the key role it plays in quantum mechanical measurement—according to PQT.

(7): If a system S_0, in the form of a spatially spread out wave packet, interacts inelastically and simultaneously with N highly localized systems $S_1 \ldots S_N$ ($N \geqslant 2$), in such a way that for each interaction the conditions for a probabilistic event to occur are satisfied (depending on which of (6A) to (6E) turns out to be correct) then *either* S_0 interacts with just *one* of $S_1 \ldots S_N$ in a highly *localized* way (and there is a probabilistic collapse of the wave packet of

S_0) *or* S_0 interacts elastically with *all* $S_1 \ldots S_N$, interference may result, and there is no reduction of the wave packet. If the state of S_0 just before the interaction is ψ_{S_0}, then the probability that S_0 interacts in a localizing way with S_r so as to be localized within dV_r, is given by $|\psi_{S_0}|^2 dV_r$, assuming here that the probability of an elastic interaction with $S_1 + S_2 + \ldots S_N$ equals zero. Having been localized by the system S_r in this way, S_0 may of course go on to interact with further systems—as when electrons leave trails in cloud or bubble chambers, or in photographic plates.

It deserves to be noted that the phenomenon of coherent inelastic diffraction—of neutrons by a crystal, for example—does not refute (7). For interference to occur, it is necessary that the quantum of energy taken up by the crystal is *not* localized more or less permanently at one or other molecule of the crystal, but is in the form of an unlocalized phonon. This means that, for interference to occur, neutrons must *not* interact with one or other molecule of the crystal in a way which satisfies any of (6A) to (6E)—the conditions for a localizing probabilistic event to occur.

As we shall see, (7) enables PQT to recover the predictions of OQT concerning measurement of position—and of other observables as well. Two views may be taken concerning (7). On the one hand it may be regarded as a consequence of any of (6A) to (6E). On the other hand, (7) may be regarded as a rival to (6A) to (6E), (6F) let us say, essentially a modified form of (6E). (6F) just asserts that if S_0 interacts inelastically with $S_1 \ldots S_N$ so that each interaction is in a superposition of different channel outcomes, and there is no oscillation of these states between $S_1 \ldots S_N$, then *either* S_0 interacts inelastically in a highly localized way, with one or other of $S_1 \ldots S_N$, *or* it interacts *elastically* (or inelastically with oscillations in the energy states of $S_1 \ldots S_N$) with the entire system $S_1 + \ldots + S_N$.

Since I first argued in 1972 and 1973 that QT needs to be modified so that it contains precise, elementary quantum theoretic conditions for probabilistic events to occur, all reference to measurement, observables and classical physics being eliminated from the theory (see Maxwell 1972b, 1976a, Jammer 1974, pp. 520–1), a few people have, independently, sought to modify QT in this way by means of postulates different from those considered above. There is the postulate of Bedford and Wang (1975, 1977) which asserts that if two systems A and B interact briefly to form the superposition: $c_1|A_1\rangle|B_1\rangle + c_2|A_2\rangle|B_2\rangle$ with $\Delta E = |E_{A_1} - E_{A_2}| = |E_{B_1} - E_{B_2}|$, where E_{A_1}, etc. are the energies of the systems, then the superposition collapses spontaneously into *either* $|A_1\rangle|B_1\rangle$ or $|A_2\rangle|B_2\rangle$ with probabilities $|c_1|^2$ and $|c_2|^2$ respectively. Following up the idea of Bedford and Wang, there is Bussey's proposal (1984, 1986) that elastic collision leads *either* to collision *or* to non-interaction—an idea subsequently extended to include decay (Bussey 1987). There is the proposal of Ghirardi et al. (1986) according to which wave packets spontaneously localize in such a way that this occurs infrequently for the isolated particle but frequently for many particle systems. Finally, there is Penrose's suggestion (1985) that it is *gravity* which induces superpositions to decay into one or other state: see also Károlyházy (1966).

I have now four comments to make about the six possible postulates I have put forward above, each a slightly different version of the basic idea that it is particle creation which leads to probabilistic wave packet collapse.

1. These postulates may be regarded as distinct dynamic superselection principles, in that each postulate denies the existence (or persistence) of superpositions asserted to exist by OQT. This is somewhat analogous in form to generally accepted superselection principles which deny the possibility of superpositions of states of angular momenta that are an integer and half-integer multiple of $\hbar$, or of states of different total charge (Wick et al., 1952).

2. Any version of PQT must satisfy what may be called a principle of 'local-global' phase invariance. If two systems, A and B, have not yet interacted, then the phase of the state of A may be changed globally without this affecting the state of A+B. In this way, global phase invariance can be applied locally, to a part of A+B. Both OQT and PQT comply with this application of 'local-global' phase invariance. In addition PQT (but not OQT) asserts that if A and B interact and separate spatially and A interacts with C so that A+C undergoes a probabilistic transition at time t, then after t, the phase of the state of A+C can be arbitrarily changed globally, without this affecting the state of (A+C)+B.

This principle of local-global phase invariance provides a basis for counting propensitons (not at all the same as counting 'particles'). Given a composite system S, if the state of S can be factorized into no more than n states $\psi_1 \ldots \psi_n$, such that local-global phase invariance can be applied to each ψ_r at time t then, at this time, S is made up of just n distinct propensitons.

3. A decisive feature of (6A) to (6F) is that they each assert that probabilistic events occur whenever energy in the form of rest mass is converted into other forms of energy, or *vice versa*. This may seem to be a not altogether implausible postulate when it is put into historical context. From its inception, QT was developed in response to problems concerning the interaction of radiation and matter. This is true of Planck's original quantum theory of black-body radiation; it is true of Einstein's photon hypothesis of 1905, in terms of which hitherto puzzling aspects of the photoelectric effect were explained; it is true of Bohr's 1913 quantum theory of the atom, which explained the manner in which atoms absorb and emit radiation (in accordance ith the Balmer series); and it is true of subsequent contributions to the development of QT made by Heisenberg, Schrödinger, Dirac and others (Jammer, 1966). If QT is understood in this way, as concerned quite fundamentally, and from the outset, to solve problems concerning the interaction between matter and radiation, then

the postulate that probabilistic events occur when rest mass is converted into or created out of radiant (or other) forms of energy, can be seen to be a rather natural development of QT, in no way arbitrary or *ad hoc*.

4. One feature of PQT may be held to be problematic. PQT postulates *instantaneous* collapse of wave packets, as a real physical phenomenon. (The collapse is instantaneous in the centre of mass frame of the system that becomes 'actualized' with the occurrence of the probabilistic event.) Such instantaneous collapse conflicts with special relativity.

This feature of PQT is not however as damaging as it may at first sight appear to be. Five points deserve to be made.

First, OQT (in its relativistic versions) evades this conflict with special relativity—insofar as it does—only by being extremely *imprecise* as to what does occur when measurements are made. It is this lack of precision which makes it possible to hold, within the framework of the orthodox viewpoint, that the collapse of the wave packet, associated with measurement, is not a real physical process, and therefore not a process whose instantaneous occurrence can conflict with special relativity (SR). PQT, we may say, as a result of being much more precise than OQT, makes explicit a non-relativistic feature of quantum theory that is only implicit in the vaguer, more ambiguous OQT. It deserves to be noted that attempts that have been made to provide a more precise, Lorentz invariant theory of wave packet collapse within the framework of OQT have not met with success (Aharonov and Albert, 1981).

Second, granted that wave packet collapse is a real physical phenomenon, recent experimental results of Aspect et al. (1982) reveal that this phenomenon occurs in a faster-than-light way. To this extent, the non-relativistic collapse of the wave packet, predicted by PQT, has been experimentally corroborated. (Aspect et al., however, interpret their result differently, as confirming the anti-realist feature of OQT.)

Third, I have shown elsewhere (Maxwell, 1985) that probabilism in general is incompatible with SR. But if this is the case, then it can be no defect whatsoever of probabilistic PQT that it is incompatible with SR. Once again, we see that OQT only evades this incompatibility because of its ambiguity concerning the crucial issue of whether or not the quantum domain is fundamentally probabilistic in character.

Fourth, PQT is incompatible with SR in only an extremely subtle way. It is only the manner in which physical potentialities evaporate (as one may put it) with the occurrence of probabilistic events, that contradicts SR. The dynamical evolution of quantum systems is otherwise, according to (relativistic) PQT, fully Lorentz invariant. If SR is interpreted phenomenalistically, as prohibiting instantaneous *signals*, then SR and PQT become compatible insofar as instantaneous wave packet collapse cannot be used to transmit signals.

Fifth, the subtle way in which PQT conflicts with SR appears to be related to the dramatic way in which QT more generally conflicts with general relativity (GR) in connection with Hawking radiation of black holes. Both kinds of conflict arise in connection with the conversion of rest mass into radiant energy.

GR implies that a non-rotating black hole cannot lose mass. Hawking (1974) showed that QT implies that at the event horizon of a black hole, the gravitational field causes photon pair creation in such a way that one photon of each pair escapes to infinity, thereby decreasing the rest mass of the black hole. This quantum process leads to the eventual 'evaporation' of the black hole, in sharp conflict with GR.

Once it is conceded that the conversion of rest mass into radiant energy can (in certain circumstances) lead QT to contradict GR, it is perhaps not so very implausible to suppose that the same physical process can lead QT to contradict SR (in a fashion which experiment corroborates and does not refute). That PQT does subtly contradict SR in this way may well indicate that PQT contains important clues as to how QT and GR are to be unified.

8 EXPERIMENTAL SUCCESS OF OQT ENSURES EXPERIMENTAL SUCCESS OF PQT

In what follows, PQT is to be understood as the dynamical equations of QT interpreted and restricted in scope by postulate (5) and one or other of (6A) to (6F), understood as different more precise versions of (2B).

Putting on one side for the moment the few special conceivable experiments for which OQT and PQT give slightly different predictions, I now show that PQT recaptures all the experimental success of OQT, even though PQT (like Newtonian mechanics and Maxwellian electrodynamics) makes no reference to observables or measurement in its basic postulates.

It is, to begin with, not hard to see how PQT can reproduce the experimental predictions of OQT as far as *position* measurements are concerned at least. A quantum position measurement invariably involves the detection of a quantum object by means of the occurrence of an inelastic collision, which creates a new particle state capable of providing a permanent, detectable record of what has occurred. Thus, for example, an atom is ionized, creating a detectable dot in a cloud or bubble chamber; a silver bromide molecule is dissociated, creating a detectable dot of silver in a photographic emulsion after development. Analogous remarks hold for scintillation and geiger counters. All such processes, associated with position measurements, with the detection of quantum objects, are precisely of the kind that are, according to PQT, associated with the probabilistic actualization of propensities.

It becomes clear that PQT is able to predict the results of measuring all other observables besides position once one realizes that measurements of all other observables invariably involve position measurements, the detection of quantum objects, of the kind just indicated. It is of decisive importance, here, to distinguish *preparation* and *measurement*. Preparation arranges for quantum objects to have some definite quantum state in some specific spatial region: quantum objects are not detected, and no probabilistic events need occur. Measurement, on the other hand, invariably involves the detection of quantum objects—and thus processes of the kind just indicated. A measurement of spin, momentum or energy typically involves first, a preparation procedure to ensure that eigenstates of the observable can be associated with distinct spatial regions, and then, second, a position measurement to detect the quantum object in one or other region. PQT, in predicting the outcome of inelastic collisions that create new particle, stationary states, is thus able to predict the results of measuring *all* quantum observables (via the detection of quantum objects, the measurement of position). PQT, indeed, does much better justice to the realities of quantum measurement than OQT (yet another indication of its superiority). Whereas the formalism of OQT suggests that all observables are on an equal footing, PQT makes it quite clear that this is not the case at all.

Finally, we have every reason to hold that PQT recaptures in detail *all* the experimental success of OQT, even though PQT treats measurement in a purely quantum mechanical way, and makes no use of classical physics for a specification of physical states of measuring instruments, in the manner of OQT. Because OQT and PQT share all the same dynamical equations, *the very empirical success of OQT itself ensures this result.* If physical processes associated with quantum measurements cannot be accurately predicted (in principle at least) by quantum scattering theory, as developed in the literature, but reinterpreted in terms of PQT, then this would imply that OQT itself is empirically defective.

Essentially only two reasons exist for holding that orthodox quantum collision theory does not apply to measurement. First, probabilistic events that involve wave packet collapse are it seems to be associated with measurement: no such events are to be associated with quantum wave packed collisions, according to orthodox quantum collision theory. Second, OQT is interpreted to be about the results of performing measurements on quantum systems: hence OQT cannot, without incoherence, be applied to the process of measurement itself. PQT solves both these problems of principle. First, according to PQT, inelastic wave packet collisions do indeed quite generally involve probabilistic events and wave packet collapses of precisely the kind to be associated with measurement. Second, PQT reinterprets the ψ-function to contain probabilistic information about the outcome of inelastic wave packet collisions: all measurements are just special cases of such inelastic wave packet collisions. Thus the decisive objections to applying *orthodox* collision theory to measurement do not arise when *propensity* collision theory is applied to measurement. In moving from OQT to PQT, (i) quantum wave packet collision theory, and (ii) the theory and role of measurement in quantum theory, are *both* adjusted, so that the former becomes straightforwardly applicable to the latter. Formidable technical problems may well arise, having to do with solving Schrödinger's equation for many interacting systems; but problems of principle do not arise. (PQT can of course avail itself of the orthodox account of measurement as a matter of practical convenience, in order to simplify calculations: the vital point is that PQT, unlike OQT, does not need to do this as a matter of conceptual necessity.)

9 CRUCIAL EXPERIMENTS

The six versions of PQT put forward here are by no means equally easy to distinguish from OQT on experimental grounds. The version that departs most radically from OQT on experimental grounds is PQT based on (6A). It may well be that existing inelastic wave packet scattering data already refute (6A).

At the opposite extreme, PQT based on (6F) may well be in practice (even if not in principle) indistinguishable experimentally from OQT.

Two kinds of crucial experiment are of decisive importance. The first has to do with potential particle interference, the second with decaying systems—with the rates at which such systems decay, and possible deviations from exponential rates of decay.

The first kind of experiment can be illustrated by the following thought experiment. A particle, about to decay into two equal fragments, moves towards a three-slitted screen. The particle has a certain propensity to decay before encountering the screen, at region A, and a certain propensity to decay after passing through the screen, at region B. According to OQT, the particle passes through the slits of the screen as a *superposition* of the undecayed state (which passes through the central slit) and the decayed state (which passes through the two outer slits in the form of two decay fragments). PQT based on (6A) asserts that this superposition does not exist if time of flight from A to B $\Delta t > \hbar/\Delta E$, since in this case the superposition of undecayed and decayed states jumps probabilistically into one or other state. If the half life of the decay is comparable to the time of flight from A to B, then PQT based on (6D) predicts that the superposition does not exist in half of an ensemble of similarly prepared systems. Let subsequent position measurements of decay fragments (by means of a photographic film) be such that decay at A or B are indistinguishable. In this case interference is possible, since as the relative distance between the film and A and B vary, so wave packets of decay fragments interfere constructively or destructively, in this way creating interference bands on the film when the experiment is repeated many times. OQT predicts that such interference bands exist whereas versions of PQT predict that they do not or only exist weakly.

Another version of this first kind of experiment would involve arranging for a photon to encounter a nucleus in such a way that there is a probability $=1/2$ that the photon persists, probability $=1/2$ that an electron/positron pair is created. The experimental arrangement is such that these two channels are then reflected back to the region of the interaction in such a way that interference effects arise if *both* channels exist. In such circumstances OQT predicts interference whereas appropriate versions of PQT predict no (or weak) interference.

A second kind of crucial experiment exploits two different ways in which, according to OQT, any quasi-stable system can decay. Type 1 decay occurs when the conditions are such that nothing exists to subject the system to measurement: in this case, according to OQT, the system persists as a superposition of the undecayed and decayed states until a measurement is eventually performed. Type 2 decay occurs when measurements capable of detecting whether or not the system has decayed, are performed at intervals Δt; in this case, according to OQT, the system returns to its initial undecayed state every Δt until the system is eventually detected to have decayed.

OQT predicts that type 1 and 2 decays proceed at very nearly the same rates. But not quite. For times considerably longer than Δt, type 2 decay is invariably exponential (Maxwell, 1973). Type 1 decay must however, according to OQT, depart slightly from the exponential for long times (Fonda et al., 1978). This difference makes the second kind of crucial experiment possible. Leaving (6C) on one side, the remaining five versions of PQT predict that, in appropriate circumstances, systems decay in a type 2 way when OQT predicts type 1 decay. This difference arises whenever conditions are such that probabilistic events occur, according to the relevant version of PQT, even though no measurement is performed. In such circumstances, for long times, the relevant version of PQT predicts exponential decay, whereas OQT predicts slight departure from exponential decay. Accurate determination of whether or not decay is exponential for long times in various physical conditions can therefore decide between OQT and at least five versions of PQT. Attempts have been made to verify non-exponential decay, but the experimental results so far appear to be inconclusive: see Butt and Wilson (1972).

OQT predicts also that type 1 decay is non-exponential for short times (Fonda et al., 1978). This carries with it the implication that rapid repeated measurement can effect the rate of decay. There is here the possibility of another kind of experiment capable of distinguishing between OQT and versions of PQT: for a discussion of the theoretical and experimental issues involved here, see Fonda et al. (1978) and Ghirardi et al. (1979).

It deserves to be noted that OQT already faces one serious problem in connection with decay: OQT implies that a continuously observed system cannot decay at all! For a discussion of this problem, and how OQT may be modified so as to overcome it, see Sudbery (1984).

IO WHY HAS QUANTUM PROPENSITON THEORY BEEN OVERLOOKED?

Whereas OQT fails to solve the problem of what sort of entities quantum objects can be in virtue of their ostensibly contradictory wave-like and particle-like properties, PQT provides a precise, consistent solution to the problem: quantum objects are varieties of a distinctively quantum mechanical kind of *discrete propensiton.* Granted that the quantum world is fundamentally probabilistic (a reasonable hypothesis) this characterization of quantum objects as discrete propensitons arises in wholly natural way. Whereas OQT is grossly *ad hoc,* vague, ambiguous and non-explanatory, PQT is, in comparison, non-*ad hoc,* powerfully explanatory and precise (even granted the uncertainty as to which of (6A) to (6F) should be adopted). PQT is free of the seven very grave difficulties that plague OQT. Furthermore, crucial experiments appear to be possible capable of deciding between OQT and versions of PQT. The case for taking PQT sufficiently seriously to develop it further theoretically, and put it to the test experimentally, would seem to be overwhelming. Why, then, has this propensiton approach been ignored for the last sixty years, despite the flood of literature on problems concerning the interpretation of QT?

A part of the answer to this question must be that the propensiton approach advocated here has not been entirely overlooked. Thus de Broglie, Vigier, Landé and Popper have all sought to interpret QT in ways which combine realism and probabilism—the key idea of the propensiton approach. Unfortunately these thinkers have persisted in the attempt to understand quantum objects in terms of classical objects and properties, thus violating a key tenet of this paper—assumption (I) of section 3. Thus de Broglie's theory of the double solution conceives of quantum objects as *particles* guided by pilot *waves* (de Broglie, 1964). Landé (1965) holds quantum objects to be *particles.* And so does Popper, despite having done more than anyone to introduce the propensity idea into quantum physics. This is in part because, for Popper, propensities are *relational* properties rather than intrinsic and fundamental properties of quantum objects *per se.* As he has put it '*Propensities are properties of neither particles nor photons nor electrons nor pennies.* They are *properties of the repeatable experimental arrangement.*' (Popper, 1967). The result, as Feyerabend (1968) has argued, is that Popper ends up, despite his intentions to the contrary, defending a position not so very different from Bohr's.

Scattered throughout the literature there are remarks—by Born, Heisenberg, Dirac, Eddington, Jeans, Landé (see Popper, 1982, pp. 130–35) and Margenau (1954)—that can be regarded as anticipating the propensity viewpoint proposed here. These remarks all fail, however, to take seriously the need to specify micro realistic conditions for probabilistic events to occur. They thus indicate variants of OQT, rather than PQT.

Ironically enough, elsewhere in the literature, in connection with decay, a number of authors express the view that measurement-type interactions occur even when only purely quantum, micro phenomena are involved: see Fonda et al. (1978, p. 623). In a recent paper Sudbery remarks that it is 'universally' assumed 'in textbooks and research papers alike' that decaying systems decay at some definite time even when not subject to measurement (Sudbery, 1984, p. 529). In other words, the key physical postulate of PQT—that inelastic interactions involve probabilistic transitions—is widely taken for granted by physicists as far as decay processes are concerned, even though this clashes with OQT. One is almost inclined to say that the theory that physicists employ in practice is PQT obscured by the rhetoric of OQT.

All this only deepens the mystery as to why PQT was not explicitly advocated long ago, soon after Schrödinger's epoch making work. The solution to the mystery can, in outline, be put like this.

If, around 1926, the physics community had conceived of physics in terms of *conjectural elemental essentialism*, and had fully appreciated the point that there is an intimate link between the nature of physical objects and properties on the one hand, and the nature of dynamical laws on the other hand (assumption (I) of section 3), then PQT would almost certainly have been put forward soon after Schrödinger's work—perhaps by Schrödinger himself. This would have occurred as soon as the fundamentally probabilistic character of the quantum world had been recognized. Unfortunately, very different instrumentalist and positivist philosophies of physics prevailed at the time. Bad philosophy, in short, prevented physicists from developing PQT soon after 1926; subsequently, orthodoxy hardened into an almost uncriticizable dogma.

In a little more detail, the failure of the physics community to consider PQT during the years 1926–1935 can be explained in the following way. In order to entertain PQT as a viable possibility, it is absolutely essential to distinguish sharply between abandoning *determinism* and abandoning *micro realism* as one moves from the classical to the quantum domain. It is only if we decisively distinguish these two issues that we are able to consider *abandoning* determinism but *retaining* micro realism—which is what PQT presupposes. Quite disastrously, everyone during the years 1926–1935 conflated these two distinct issues. On the one hand, Bohr, Heisenberg and Born argued for the abandonment of determinism-and-micro-realism; on the other hand Einstein and Schrödinger argued for the retention of determinism-and-micro-realism (Jammer, 1974). Because of the universal conflation of these two quite distinct issues, no one was able to argue for the abandonment of determinism and the retention of micro realism—for the development of probabilistic micro realism, in other words. Gradually around 1934/35 Einstein and Popper did discover how to distinguish these two very different issues. But by then it was too late: the Copenhagen position had become an unassailable dogma.

A part of the reason for the general failure to distinguish 'determinism versus probabilism' from 'micro realism versus instrumentalism' was this. Everyone at the time tended unthinkingly to take subjectivist interpretations of probability for granted. Both the Bohr camp and the Einstein camp tended to assume that a probabilistic physical theory could not be about *reality*, but could only be about *our incomplete knowledge of reality*. Thus abandonment of determinism seemed to everyone to carry with it the implication that micro realism must be abandoned as well. Einstein and Schrödinger, reluctant to abandon realism, felt forced to hold on to determinism as well; Bohr and Heisenberg, seeing forcefully the need to abandon determinism, felt compelled to abandon realism as well. Thus no one was able to envisage the obvious option: probabilistic micro realism, or the *discrete propensiton!*

There is, however, a deeper reason for the general failure of the physics community to consider this possibility of a new kind of probabilistic physical entity. Again and again in the history of physics we encounter the failure of physicists to conceive of new physical objects and properties appropriate to new physical theories.

Before Newton, the new natural philosophy had been closely associated with the attempt to understand Nature in terms of corpuscles interacting by contact (Dijksterhuis, 1961; Burtt, 1932). A physical theory, in order to be an acceptable and comprehensible contribution to physics, had to be capable of being interpreted in terms of this corpuscular idea. Thus, when Newton's law of gravitation appeared, natural philosophers did not set out to invent a new kind of object, with new kinds of physical properties, appropriate to the new law. Quite to the contrary, many of Newton's contemporaries in effect took it for granted that the law could only be acceptable if explicable in corpuscular terms. Insofar as they judged this to be impossible, they found Newton's law to be incomprehensible, or without explanatory power. Thus Huygens, in a letter to Leibniz, writes: 'Concerning the Cause of the flux given by M. Newton, I am by no means satisfied [by it], nor by all the other Theories that he builds upon his Principle of Attraction, which to me seems absurd ... I have often wondered how he could have given himself all the trouble of making such a number of investigations and difficult calculations that have no other foundation than this very principle' (Koyré, 1965, pp. 117–8). In a sense, Newton himself agreed, as is indicated by his remark: 'That gravity should be innate, inherent, and essential to matter, so that one body may act upon another, at a distance through a vacuum, without the mediation of anything else ... is to me so great an absurdity, that I believe no man who has in philosophical matters a competent faculty of thinking can ever fall into it' (Burtt, 1932, pp. 265–6). The manifest impossibility of interpreting the law of gravitation in corpuscular terms led Newton to adopt an instrumentalist or positivist interpretation of the law, according to which the law merely describes how objects move, without specifying the cause of such motions, or providing any kind of explanation for the motions. In other words, the failure of an utterly misguided attempt at a realist or essentialist interpretation of the law, led straight to instrumentalism and positivism.

Subsequently, however, Boscovich and others were able to conceive of an object with properties rather more appropriate to Newtonian theory—namely, the point-particle, possessing inertial mass, and being surrounded by a spherically symmetrical, centrally directed force-field, varying continuously with distance, but otherwise invariant. This genuinely Newtonian object can be regarded as a generalization of the pre-Newtonian solid corpuscle (retained by Newton himself). It accords beautifully with Newton's law of gravitation, having been designed in part just for that purpose.

This pattern of confusion arises again in the late nineteenth and early twentieth century, in connection with Maxwell's new theory of electromagnetism. Instead of trying to invent a new kind of object—the electromagnetic field—appropriate to Maxwell's equations of electromagnetism, Maxwell himself, Kelvin and others devoted endless effort to the attempt to understand the new theory in terms of old, pre-Maxwellian objects—namely, the material aether, in turn to be understood, presumably, in terms of Boscovichian point-atoms, or even pre-Boscovichian corpuscles. The failure—indeed the frequent absurdity—of these (utterly misguided) attempts at a realist or essentialist interpretation of electromagnetism led many to adopt instrumentalist and positivist viewpoints.

It was only when special relativity became accepted that physicists began to appreciate what Faraday had seen all along—namely that the electromagnetic field is a new kind of object with new properties (a generalization of the Boscovichian force-field), which does not need to be understood in terms of old, inappropriate particle-like objects and properties. Indeed, instead of trying to understand the electromagnetic field in terms of some matter-like substance, one should rather try to understand matter in terms of the field.

The lesson from history is I hope clear. The same pattern of confusion arises in connection with quantum theory. Instead of trying to invent a new kind of physical object (the discrete propensiton!) appropriate to the character of the new theory, physicists rather persisted in the attempt to interpret the new theory in terms of old, inappropriate, deterministic objects—the point-particle and the field—lapsing into instrumentalism and positivism when this utterly misguided attempt failed.

What is needed, in order to avoid further repetitions of this pattern of confusion is the general adoption of what I have called *conjectural elemental essentialism*, which carries with it the implication that to talk of theoretical objects and their properties *is* to talk of the dynamical laws that the objects obey (a change in our ideas of the latter thus automatically necessitating a change in our ideas of the former). In this way we may in future avoid what is so striking a feature of the history of theoretical physics: great imaginativeness and inventiveness concerning equations and theories, combined with immense conservatism and stupidity concerning physical *objects* and their *properties*. (More fundamentally, we need the adoption of what I have called elsewhere *aim oriented empiricism*, to be implemented within the general framework of the *philosophy of wisdom*: see Maxwell, 1984.)

It deserves to be noted that the one great exception to the above pattern of confusion is Einstein's invention of general relativity: here, from the outset, there is harmony and accord between the new theory, and the new object the theory is taken to describe—curved Riemannian space-time. (But this exception is hardly surprising, since Einstein came close, instinctively, to pursuing natural philosophy in accordance with aim oriented empiricism and the philosophy of wisdom.)

II CONCLUSION

Two and a half thousand years ago the presocratic philosophers first tried to explain and understand the world in terms of some kind of elemental stuff, invariant through all change and diversity. For Anaximander, everything was diverse, lawful manifestations of the *apeiron* or boundless; for Heraclitus everything was lawfully regulated fire or process; for Democritus everything was the outcome of intrinsically unchanging atoms in relative motion in the void. After Plato and Aristotle, the astonishing endeavour of the presocratics fell into decay, and was abandoned. It was resurrected in the seventeenth century by those who created modern science: Galileo, Kepler, Descartes, Huygens, Boyle, Newton. From that time until the present, the basic idea of the presocratics has been the fundamental guiding idea of physical science: to explain and understand change and diversity in terms of that which is elemental and invariant. But it is above all in the twentieth century, and especially in the last decade or so, that giant strides have been made towards fulfilling the presocratic vision. Grand unified theories, quantum gravity and superstring theory are, for the first time ever, groping attempts at a unified scientific theory of everything. And here we come to the paradox. For just when the two and a half thousand year old research programme of the presocratics seems close to completion, the physics community has lost interest. The orthodox version of QT is, as we have seen, hopelessly inadequate from the standpoint of enabling us to explain and understand complex macro phenomena in terms of elemental micro phenomena. Insofar as the physics community, by and large, accepts OQT as unproblematic, *and does not actively seek a better alternative*, it has lost interest in the noble quest of the presocratics. As Einstein realized with anguish, the soul of natural philosophy has been betrayed. The quest to understand has disintegrated into expert puzzle solving, the hunt for Nobel prizes and defence contracts.

What I have tried to do in this paper is point to an alternative approach to understanding the quantum world which keeps alive, and does not betray, the presocratic endeavour. My hope is that others, better qualified than myself, will take up and develop the basic idea, so that it may be formulated with precision and put to the test of experiment.

12 APPENDIX

Our inability to conceive of quantum objects in classical terms—as classical particles, waves or fields—has led many to conclude that there is something inherently inexplicable about the nature of quantum objects which must prevent us from developing a fully realistic version of QT with its own consistent, self-contained quantum ontology. A central claim of this paper is that precisely the reverse of this is true. In order for quantum objects to be fully explicable and understandable it is absolutely essential that they differ radically from all classical objects—granted merely that the quantum world is fundamentally probabilistic in character. It is being understandable in classical terms—as classical particles or waves—which would render probabilistic quantum objects utterly inexplicable and incomprehensible!

Quantum objects are, I have argued, unproblematic varieties of the *discrete propensiton.* In order for this solution to the wave/particle problem to be satisfactory, I need to show that the *discrete propensiton* as I have characterized it in section 3 arises in an entirely natural, non-arbitrary way as we generalize the notion of physical object and property in moving from determinism to probabilism. I begin with a characterization of deterministic properties and objects in the following eight points (see also Maxwell 1968, 1976a, 1982, 1985).

(1) Any classical, deterministic physical property—such as rigidity, elasticity, gravitational charge, electric charge, inflammability or opacity—determines how something *changes* (or resists change) in certain circumstances. Thus, if an object is inflammable then, of necessity, it bursts into flames when the relevant physical conditions are satisfied (exposure to a naked flame, etc.). If it does not burst into flames in these circumstances then, *ipso facto*, it is not inflammable. Again, if two objects possess the property of Newtonian gravitational charge, g_1 and g_2, (equal to inertial masses m_1 and m_2), then of necessity, in the absence of other forces, the two objects accelerate towards each other in accordance with $F = Gg_1g_2/d^2$ and $F = ma$. If they do not, then *ipso facto*, they do not possess Newtonian gravitational charge. Yet again, the classical electromagnetic field can be regarded as having, at any given point and instant, a value of electric and magnetic field intensity, which determines the way a test particle would accelerate were it to occupy the given point at the given time. In this way, the physical properties of the electromagnetic field determine necessarily how charged test particles change their state of motion.

(2) Properties of the kind indicated determine how objects change (in

appropriate circumstances) with *necessity*, as long as the relevant property is possessed by the object in question. David Hume (1959) is wrong. It *is* possible for there to be necessary connections between successive states of affairs (Maxwell, 1968). All that we require is a state of affairs made up of objects with invariant, necessitating, deterministic properties (of the kind just indicated). We cannot *know* (for certain) that necessary connections between successive states of affairs really exist, just because we can never *know* that the relevant necessitating properties really exist. But equally, we cannot *know* that necessary connections, and necessitating properties, do not exist.

(3) Classical theories, when understood in the usual way, do not attribute necessitating properties to physical objects. Rather they assert that certain physical objects obey *contingent regularities*. Thus Newton's law of gravitation, as ordinarily understood, does not attribute the necessitating property of gravitational charge to particles; rather it asserts that particles obey the law-like but *contingent* regularity $F = Gm_1m_2/d^2$ (and $F = ma$.). Classical theories can however be reinterpreted so that they do attribute necessitating properties to physical objects. Thus Newton's law of gravitation can be reinterpreted *essentialistically*, to assert: 'All particles have (Newtonian) gravitational charge $g = m$ (where m is the inertial mass of the particle)'. Here, it is built into the *meaning* of 'Newtonian gravitational charge g' that if bodies possess this property then, of necessity, they obey Newton's law $F = Gm_1m_2/d^2$ (and $F = ma$.). The outcome of interpreting Newtonian theory in this way is that the law becomes an analytic statement—a statement true in virtue of the meaning of the constituent terms. And quite generally, interpreting a classical theory *essentialistically*, so that it can be used to attribute necessitating properties to physical objects, renders all the law-like statements of the theory *analytic statements*. But no loss of empirical content is involved in interpreting physical theories in this way—a crucial point to appreciate. The whole empirical content of an essentialistically interpreted theory is contained in the assertion that all physical objects (of the specified type) do in fact possess the relevant physical properties, in virtue of which, according to the theory, the laws are of necessity obeyed. Thus the whole empirical content of Newton's theory of gravitation, essentialistically interpreted, is concentrated in the statement that all objects possess Newtonian gravitational charge, equal in value to inertial mass. It is *this* statement that is refuted with the discovery of objects moving under the influence of gravitation in a way which violates Newton's law. (In a sense, general relativity, if true, does not refute the essentialistic version of Newton's law of gravitation: rather, it shows that strictly nothing exists to which Newton's law is applicable.)

(4) Essentialistically interpreted physical theories provide an *explanation* for the existence of the lawful regularities they postulate. The regularities exist because physical properties exist which make such regularities inevitable. If essentialistic Newtonian theory is true, then all objects really do possess

Newtonian gravitational charge, and it is inevitable, necessary, that all objects obey Newton's law of gravitation. By contrast, non-essentialistic theories can provide no explanation as to why postulated lawful regularities are observed. Theories so interpreted are incapable of postulating the *existence of anything responsible for* lawful regularities. As restricted regularities are derived from more and more comprehensive or universal ones, so the mystery as to why such regularities should obtain in nature at all can only deepen.

(The point made here—that essentialistic theories have greater explanatory power than non-essentialistic theories—is of course wholly in addition to, and independent of, the point made earlier, that conceptually coherent, *non-ad hoc* theories are, other things being equal, more explanatory than incoherent, *ad hoc* theories, such as OQT.)

(5) The greater explanatory power of essentialistic theories (of the kind just indicated) makes it desirable to commit theoretical physics as a whole to the aim: to discover those few basic invariant essentialistic properties (possessed by a few different sorts of basic physical entities) which we conjecture to exist and to be responsible for all change and diversity in the world. The view that physics ought to have this aim of discovering a true, unified, comprehensive, *essentialistic* physical theory, unifying all forces and applicable to all phenomena, may be called *conjectural elemental essentialism*. This paper in effect seeks to indicate one way in which OQT can be modified so that it accords rather better with conjectural elemental essentialism.

(6) Classical deterministic physical properties have associated with them *two kinds of change*. First, there is the change which occurs when the property is 'actualized': the inflammable wood burns, the elastic ball bounces, the gravitationally charged objects accelerate towards each other. Second, values of the property may themselves change: the inflammable wood may gradually become less inflammable (as it becomes wet), the ball may become less elastic. Newtonian theory postulates that gravitational charge is *invariant*: and it is of course the task of theoretical physics to discover properties that are invariant through as wide a range of changes as possible. There is always the possibility, however, that gravitational charge changes (as indeed it does given relativity).

(7) Classical deterministic physical properties are of two kinds. The first kind are properties which are actualized discontinuously or *discretely*, when special physical conditions arise, such as when the inflammable object is exposed to a naked flame, or elastic objects collide. The second kind are properties which are actualized *continuously*, as in the case of Newtonian gravitational charge (granted that there is more than one object in the universe).

(8) The nature or character of any classical deterministic *object* is entirely given or determined by the nature of the *properties* possessed by the object. There is nothing to the object over and above the properties it possesses. The nature of the Newtonian point-particle or Maxwellian electromagnetic field is

entirely specified by the physical properties these objects possess—and hence by the dynamical laws these objects obey.

All these points concerning classical deterministic properties and objects carry over to fundamentally probabilistic properties and objects, to *propensities* and *propensitons.* The *only* difference is that when a deterministic property is actualized there is just one possible outcome, whereas when a propensity is actualized there are n possible outcomes. A specific *value* of a propensity P specifies n probabilities $p_1 \ldots p_n$, and attributes a definite probability p_r to each possible outcome O_r, with $\sum_{r=1}^{n} p_r = 1$.

Just as classical, deterministic properties determine how things change (point (1)), so too propensities determine how things change in certain circumstances, but probabilistically and not deterministically. Again, just as there can be necessary causal connections between successive states of affairs, given deterministic (essentialistic) properties (point (2)), so too there can be *probabilistic* necessary causal connections between successive states of affairs given that propensities and propensitons exist. Yet again, just as classical deterministic properties can change in two kinds of ways (point (6)), so too propensities (and propensitons) can change in two kinds of ways: the propensity can be actualized with the occurrence of a probabilistic event, and the value of the propensity itself can change. (In the case of a die, this second kind of change would take place if the die, itself a magnet, moves through a varying magnetic field.)

Finally, and most important, just as classical deterministic properties are of two kinds (point (7)), so too are propensities (and propensitons). On the one hand, corresponding to inflammability or rigidity, there are discrete propensities (and discrete propensitons), actualized probabilistically at *discrete* times, when appropriate physical (propensiton) conditions arise. On the other hand, corresponding to Newtonian gravitational charge, there are continuous propensities (and propensitons) actualized continuously in time.

Corresponding to these two kinds of propensities and propensitons (derived from the two kinds of classical properties and objects), there are two kinds of fundamentally probabilistic dynamical theories, namely *discretely probabilistic* theories (which assert that values of propensities evolve deterministically until the physical conditions arise for a probabilistic actualization to occur, when an instantaneous probabilistic event occurs, determined probabilistically by the values of the relevant propensities at that instant), and *continuously probabilistic* theories (which assert that probabilistic events occur continuously in time).

The vital point to appreciate is that the three kinds of dynamic theories that have been considered—deterministic, discretely probabilistic and continuously probabilistic—deserve to be regarded as equally viable from an *a priori* standpoint. Equally, the three kinds of *objects* that correspond to these three kinds of theories—deterministic object, discrete propensiton, and continuous propensiton—deserve to be regarded as equally intelligible or understandable, even if we may be more familiar with (approximately) deterministic objects. There is nothing intrinsically *ad hoc* or inexplicable about the instantaneous probabilistic transitions of discretely probabilistic theories (whether these arise in connection with QT or some other theory): such transitions are an inherent feature of this kind of theory. Furthermore, to demand that such instantaneous probabilistic transitions (e.g. those associated with PQT) must be explained in terms of some continuous evolution or theory is just to refuse to recognize the discrete propensiton, and the discretely probabilistic theory, as viable possibilities.

University College, London

REFERENCES

AHARONOV, V. and ALBERT, D. Z. [1981]: *Phys. Rev.*, **24D**, p. 359.

ASPECT, A., GRANGIER, P. and ROGER, G. [1982]: *Phys. Rev. Lett.*, **49**, p. 91.

AUDI, M. [1973]: *The Interpretation of Quantum Mechanics*, Chicago University Press.

BEDFORD, D. and WANG, D. [1975]: Towards an Objective Interpretation of Quantum Mechanics, *Nuovo Cim.*, **26B**, pp. 313–25.

BEDFORD, D. and WANG, D. [1977]: A Criterion for Spontaneous State Reduction, *Nuovo Cim.*, **B37**, pp. 55–62.

BOHM, D. and HILEY, B. J. [1987]: Non-relativistic Particle Systems, *Phys. Rep.*, **144**, pp. 321–75.

BOHR, N. [1949]: Discussion with Einstein on Epistemological Problems of Atomic Physics, in Schilpp, P. A., (ed.) *Albert Einstein: Philosopher-Scientist*, Harper, pp. 199–241.

BORN, M. [1926]: *Zeitschrift fur Physik*, **37**, p. 863.

BORN, M. [1927]: *Nature*, **119**, p. 354.

BURTT, E. A. [1932]: *The Metaphysical Foundations of Modern Physical Science*, Routledge and Kegan Paul.

BUSSEY, P. J. [1984]: When does the Wavefunction Collapse?, *Phys. Lett.*, **106A**, pp. 407–9.

BUSSEY, P. J. [1986]: Wavefunction Collapse and the Optical Theorem, *Phys. Lett.*, **118A**, pp. 377–80.

BUSSEY, P. J. [1987]: The Fate of Schrödinger's Cat, *Phys. Lett.*, **120A**, pp. 51–3.

BUTT, D. K. and WILSON, A. R. [1972]: A study of the radioactive decay law, *J. Phys.*, **A5**, p. 1248.

CRAMER, J. G. [1986]: The Transactional Interpretation of Quantum Mechanics, *Rev. Mod. Phys.*, **58**, pp. 647–87.

DE BROGLIE, L. [1964]: *The Current Interpretation of Wave Mechanics*, Elsevier.

DIJKSTERHUIS, E. J. [1961]: *The Mechanization of the World Picture*, Oxford University Press.

EINSTEIN, A. [1905]: *Ann. Phys.*, **17**, p. 132 (English translation: *Am. J. Phys.*, [1965], **33**, p. 367.

EINSTEIN, A. [1950]: Letter to Erwin Schrödinger, in Przibram, K. (ed.) [1967]: *Letters on Wave Mechanics*, Vision Press, p. 39.

EINSTEIN, A. [1951]: Letter to Michele Besso, in Speziali, P. (ed.) [1972]: *Albert Einstein Michele Besso Correspondance 1903–1955*, Hermann, p. 453.

FEYERABEND, P. [1968–69]: On a Recent Critique of Complementarity, *Phil. Sci.*, **35**, pp. 309–31; **36**, pp. 82–105.

FEYNMAN, R. P., LEIGHTON, R. B. and SANDS, M. [1965]: *The Feynman Lectures on Physics*, vol. 3, Addison-Wesley.

FINE, A. [1970]: Insolubility of the Quantum Measurement Problem, *Phys. Rev.*, **D2**, pp. 2783–7.

FONDA, L., GHIRARDI, G. C. and RIMINI, A. [1978]: Decay Theory of Unstable Quantum Systems, *Rep. Prog. Phys.*, **41**, pp. 587–631.

FRAUENFELDER, H. and HENLEY, E. M. [1974]: *Subatomic Physics*, Prentice-Hall.

GARDNER, M. [1972]: Quantum-Theoretical Realism: Popper and Einstein v. Kochen and Specker, *Brit. J. Phil. Sci.*, **23**, pp. 13–23.

GHIRARDI, G. C., OMERO, C., WEBER, T. and RIMINI, A. [1979]: Small-Time Behaviour of Quantum Nondecay and Zeno's Paradox in Quantum Mechanics, *Nuovo Cim.*, **52A**, pp. 421–41.

GHIRARDI, G. C., RIMINI, A. and WEBER, T. [1986]: Unified dynamics for microscopic and macroscopic systems, *Phys. Rev.*, **D**, pp. 470–91.

GOTTFRIED, K. and WEISSKOPF, V. F. [1986]: *Concepts of Particle Physics*, vol. 1, Clarendon Press.

HAWKING, S. W. [1974]: *Nature*, **248**, p. 30.

HEPP, K. [1972]: Quantum Theory of Measurement and Macroscopic Observables, *Helv. Phys. Acta*, **45**, pp. 237–48.

HUME, D. [1959]: *A Treatise of Human Nature*, vol. 1, Everyman.

JAMMER, M. [1966]: *The Conceptual Development of Quantum Mechanics*, McGraw-Hill.

JAMMER, M. [1974]: *The Philosophy of Quantum Mechanics*, J. Wiley.

KÁROLYHÁZY, F. [1966]: Gravitation and Quantum Mechanics of Macroscopic Objects, *Nuovo Cim.*, **A42**, pp. 390–402.

KOYRÉ, A. [1965]: *Newtonian Studies*, Chapman & Hall.

LANDAU, L. D. and LIFSCHITZ, E. M. [1958]: *Quantum Mechanics*, Pergamon Press.

LANDÉ, A. [1965]: *New Foundations of Quantum Mechanics*, Cambridge University Press.

MACHIDA, S. and NAMIKI, M. [1984]: Macroscopic Nature of Detecting Apparatus and Reduction of Wave Packet, in Kamefuchi, S. et al. (eds.) [1984] *Foundations of Quantum Mechanics*, Physical Society of Japan.

MARGENAU, H. [1954]: *Phys. Today*, **7**, p. 6.

MAXWELL, N. [1968]: Can there be Necessary Connections between Successive Events?, *Brit. J. Phil. Sci.*, **19**, pp. 1–25.

MAXWELL, N. [1972a]: A Critique of Popper's Views on Scientific Method, *Phil. Sci.*, **39**, 131–52.

MAXWELL, N. [1972b]: A New Look at the Quantum Mechanical Problem of Measurement, *Am. J. Phys.*, **40**, pp. 1431–5.

MAXWELL, N. [1973]: Alpha Particle Emission and the Orthodox Interpretation of Quantum Mechanics, *Phys. Lett.*, **43A**, pp. 29–30.

MAXWELL, N. [1974]: The Rationality of Scientific Discovery, *Phil. Sci.*, **41**, pp. 123–53 and 247–95.

MAXWELL, N. [1975]: Does the Minimal Statistical Interpretation of Quantum Mechanics Resolve the Measurement Problem?, *Method. and Sci.*, **8**, pp. 84–101.

MAXWELL, N. [1976a]: Towards a Micro Realistic Version of Quantum Mechanics, *Found. Phys.* **6**, pp. 275–92 and 661–76.

MAXWELL, N. [1976b]: *What's Wrong With Science?*, Bran's Head Books Ltd.

MAXWELL, N. [1977]: Articulating the Aims of Science, *Nature*, **265**, p. 2.

MAXWELL, N. [1979]: Induction, Simplicity and Scientific Progress, *Scientia*, **114**, pp. 629–53.

MAXWELL, N. [1980]: Science, Reason, Knowledge and Wisdom: A Critique of Specialism, *Inquiry*, **23**, pp. 19–81.

MAXWELL, N. [1982]: Instead of Particles and Fields: A Micro Realistic Quantum 'Smearon' Theory, *Found. Phys.*, **12**, pp. 607–31.

MAXWELL, N. [1984]: *From Knowledge to Widsom: A Revolution in the Aims and Methods of Science*, Basil Blackwell.

MAXWELL, N. [1985]: Are Probabilism and Special Relativity Incompatible?, *Phil. Sci.*, **52**, pp. 23–44.

PENROSE, R. [1985]: Quantum Gravity and State-vector Reductions, in C. J. Isham and R. Penrose (eds.) *Quantum Concepts of Space and Time*, Oxford University Press.

POPPER, K. R. [1957]: The Propensity Interpretation of the Calculus of Probability and the Quantum Theory, in Körner, S. (ed.) *Observation and Interpretation*, Butterworths, pp. 65–70.

POPPER, K. R. [1959a]: *The Logic of Scientific Discovery*, Hutchinson.

POPPER, K. R. [1959b]: The Propensity Interpretation of Probability, *Brit. J. Phil. Sci.*, **10**, pp. 25–42.

POPPER, K. R. [1967]: Quantum Mechanics without the 'Observer', in Bunge, M. (ed.) *Quantum Mechanics and Reality*, Springer.

POPPER, K. R. [1982]: *Quantum Theory and the Schism in Physics*, Hutchinson.

SUDBERY, A. [1984]: The Observation of Decay, *Ann. Phys.*, **157**, pp. 512–36.

VAN FRAASSEN, B. C. [1980]: *The Scientific Image*, Oxford University Press.

WICK, G. C., WIGHTMAN, A. S. and WIGNER, E. P. [1952]: *Phys. Rev.*, **88**, pp. 101–5.

WIGNER, E. P. [1967]: *Symmetries and Reflections*, Indiana University Press.

Brit. J. Phil. Sci. 39 (1988), 51–58 *Printed in Great Britain*

Research Programmes and Empirical Results

F. M. AKEROYD

Imre Lakatos [1970] stated the following:

> Within a research programme a theory can only be eliminated by a better theory, that is, by one which has excess empirical content over its predecessors, some of which is subsequently confirmed. And for this replacement of one theory by a better one, the first theory does not even have to be 'falsified' in Popper's sense of the term. Thus progress is marked by instances verifying excess content rather than by falsifying instances.

In this paper I propose to discuss a recent historical example which contradicts this viewpoint of Lakatos. In a previous paper (Akeroyd, [1986]) I discussed the work of Davy in 1810 on chlorine which led to the abandonment of Lavoisier's Oxygen Theory of Acids in the absence of a better theory: perhaps some philosophers feel at juncture the discipline of Chemistry was not mature enough for important philosophical principles to be gleaned from this episode. However the episode I propose to discuss occurred 100 years later when the protagonists were salaried academic research workers involved in a research programme funded by external sources and the achievement of Nobel prizes a realistic target.

The research programme which took as its 'hard core' the proposition that all problems relating to physiological and dietary areas had chemical solutions was, in 1910, examining the observation that the 'theoretically adequate' diet of purified fats, carbohydrates, proteins, minerals, roughage and water was inadequate for young rats unless supplemented by 15 millilitres of milk per day. Clearly some additional substance present in the milk was having a disproportionally large beneficial effect or possibly it was the proportions of known chemical compounds which were present in the milk which were critical. In England F. G. Hopkins favoured the first theory, while in the U.S.A. T. B. Osborne and L. B. Mendel favoured the second theory. In a paper originally submitted to this journal (Akeroyd, [1985]) a summary of 5 plausible explanations to account for the beneficial effect of the milk was presented, with a claim that Hopkins' classic paper in 1912 refuted four of these theories, leaving Hopkins' 'Accessory Food Factor (Vitamin)' Theory the successor by default.

Hopkins failed to isolate any previously unknown chemical compounds which could be argued to lend increased empirical content to this theory and the only positive evidence that in any way lent any such content was the claim that only 2 or 3 mls. of milk per day was an adequate dietary supplement. This claim was not substantiated when Osborne and Mendel checked Hopkins' work. Yet Osborne and Mendel abandoned their own pet theory within the research program and swung their support to the Vitamin Theory, a change of direction which effectively put them out of the running for a Nobel Prize.

Three theories refuted in Hopkins [1912] related to the possibilities that the milk contained factors which neutralised toxic elements in the artificial diet or appetite stimulating factors which counteracted the monotony of the artificial diet or factors promoting the digestibility of the artificial diet. Osborne and Mendel had earlier discarded these hypotheses on general grounds without precise experimentation. They wrote in their [1911]:

> The facts presented exclude the probability that monotony of diet is an insurmountable obstacle to nutritive success.
>
> Monotony of diet appears to have been over emphasised.
>
> We are presumably dealing with a chemical inadequacy rather than any toxicity and consequent lack of growth.
>
> The digestibility and utilisation of the artificial rations has never been shown to be unfavourable.

The theory to which Osborne and Mendel did incline held that it was the organic salt content of the milk which was the responsible factor. Disappointed with earlier experiments using artificial salt mixtures they had hit on the happy notion of using 'protein-free' milk as a source of inorganic salts which must be in a physiologically well balanced proportion. The great success of this expedient (young rats grew at normal rates for periods of 60–100 days) led them to conclude that it was the inorganic salt content of the 'protein-free' milk alone which was responsible for the beneficial effects.

It was this theory which was elegantly refuted by Hopkins in 1912. He obtained the salt mixture for the artificial diet from the ash of incinerated oats and incinerated dog biscuits, thus ensuring that it was in the correct natural proportions.

Whatever the beneficial effect of the addition of milk was to the artificial diet, it was not because of its inorganic salt content. Having eliminated four out of five plausible theories Hopkins concluded that the 'Accessory Food Factor' theory was the successor by default. The rest of the Scientific World awaited the response of Osborne and Mendel to these findings. The response was not slow in coming and was in some way possibly unexpected: both Osborne and

Mendel and another research group led by McCollum and Davis produced papers supporting the concept of 'accessory food factors' in the diet not only on the basis of new factual evidence but also on the reinterpretation of experimental data obtained by them prior to 1912.

In their paper Osborne and Mendel [1913] cited Hopkins [1912] approvingly stating:

> Our own experiments, as well of those of the other investigators mentioned, make it clear that something further than a sufficient supply of energy-yielding food material is required to promote a normal growth.

They commented that they had previously noted that a diet consisting of 18% protein, 29% starch, 25% lard, and 28% protein-free milk gave better than normal growth curve for young rats but after 60 days a decline set in. This decline could be arrested by transferring to a diet of 'milk-food' containing 60% milk powder, 12% starch and 28% lard. They observed:

> The conclusion seems inevitable, therefore, that our 'protein-free milk foods' are deficient in, or completely lack, something the milk contains and which is indispensible for perfect growth.
>
> It seems probable that the missing substance is organic in nature; for the 'protein free milk' may be presumed to contain all of the inorganic constituents of the milk.

Because of the success of the 'milk-food' which contained 16% butter, Osborne and Mendel were led to suspect that the unknown 'essential' accessory factor was present in the cream of the milk and supplied unsalted butter as a substitute for 'part of the lard' in their experimental diets. The great success of the new diet led them to conclude:

> It would seem, therefore, as if a substance exerting a marked influence on growth were present in the butter, and that this is largely, if not wholly, removed in the preparation of our natural 'protein-free milk.'

In other words Osborne and Mendel were proposing that the absence of an unisolated fat soluble vitamin was responsible for the failure of rats to survive on conventional artificial diets.

It had been important for Hopkins to feed the dietary supplement in the form of liquid milk at a separate time of day from the solid food in order to eliminate the hypothesis that the supplement acted as a flavour enhancer. However, as indicated above, Osborne and Mendel did not attach any importance to this hypothesis. They were more impressed by the observation that small amounts of dried milk powder, mixed in with solid food, were just as effective as a supplement as the liquid milk in Hopkins' experiments. From their point of view, this observation was much easier to check using their own experimental methods rather than attempt to reproduce Hopkins' exact protocols and stated experimental conditions. Their conclusion referring to 'essential' accessory factors in the diet could be deduced entirely from their own results without extending the research of Hopkins. Osborne and Mendel admitted that much of the data they had collected prior to 1912 could be interpreted as consistent with Hopkins' concept although they themselves had interpreted them otherwise. It seems quite likely that Osborne and Mendel suffered a Kuhnian-type 'Gestalt Switch': the necessary data existed prior to the crisis and the revolutionary reconceptualisation allowed them to be seen in a new way.

Once Hopkins' claim had been made and confidently asserted it was actually unnecessary to reproduce his results. It was the logic of his paper which stimulated Osborne and Mendel to look at their own previous results in a new light, and reject the inorganic salt theory. Although they decided on a crucial experiment to test the accessory food factor theory it is obvious that they were impressed by the new theory before the results of the crucial experiment described. The only 'surprising' result from the crucial experiment was the observation that the unsalted butter supplement was just as effective when added to a diet containing laboratory chemicals as a source of inorganic salts as it was when added to the diet containing 'protein—free milk' as the inorganic salt source. It was this 'unexpected' result which provided them with a result which was 'surprising' relative to their own background knowledge. With refreshing frankness Osborne and Mendel admitted their research pathway involving the concept 'protein-free milk' had been a blind alley. None of this supports the position of Lakatos with his theory of 'battles of attrition' between competing ideas, or his statement that if Nature shouts 'No', human ingenuity may always shout 'Yes' (Lakatos, [1970]).

McCollum [1913] described similar results to Osborne and Mendel but included details of the exact quantities of fat based supplement in his paper. He found that 0·5 g. of ether soluble egg yolk extract per day were required or 10% of butter in the diet. The amount of solid supplement is approximately five times the amount of fatty material present in 3 mls of milk. Neither McCollum nor O & M commented at this time on this apparent discrepancy as it was irrelevant to the main argument and they had transferred allegiance to the accessory food factor theory as a result of their own investigations. McCollum isolated the first potential vitamin and thus provided the theory with its first increase of empirical content. Osborne and Mendel were relegated to role of being chief propagandists. The great German nutritionist Rohmann later described Hopkins as being 'der Geistige Vater de Vitaminlehre' but he was not convinced of this immediately: as late as 1917 Osborne and Mendel [1917] were quoted by Lothrop [1917] as follows:

> Osborne and Mendel take exception to the recent opinion expressed by Rohmann that accessory foodstuffs are not necessary for the continued maintenance of full grown animals and that if the long familiar nutrients are available are suitable in quality and quantity, nothing further is required in the ration. They point out

> that the methods used by Rohmann to exclude the vitamines from the diets employed cannot, in any sense, be depended on to insure the absence of all traces of accessory substances. The experience of the authors and numerous other investigators has demonstrated adequate dietaries require at least two formerly unappreciated components, the so called water soluble and fat soluble accessories.

Bradford College

NOTE

McCollum [1913] wrote:

> During the past year we have been engaged in a study of composition and quantity of the inorganic content of the ration on growth in the rat . . . it should be borne in mind that these experiments were intended primarily to furnish data as to the values of the salt mixtures supplied by the rations and do not in all cases represent the highest degree of success attained by feeding the rations . . . we are convinced that these extracts contain some organic complex without which the animals cannot make further increase in body weight.

The importance of Hopkins' [1912] paper is entirely in its NEGATIVE results. He demonstrated conclusively that the utilisation of artificial rations *WAS* indeed unfavourable. In the absence of this negative proof it is most unlikely that McCollum and Davis, and Osborne and Mendel would have abandoned their pet ideas and started to invest time and effort looking for the elusive accessory food factors. Once McCollum had achieved what the original proponents of the Vitamin Theory had been unable to do, namely isolate some organic complex which acted in a catalytic or regulatory manner, then the research programme proceeded in a manner similar to that proposed by Lakatos: many older research workers and professors changed allegiance on the grounds of steadily increasing empirical content. However this is essentially a passive exercise: the required content can only be generated by research workers at the frontiers of science who have changed their ideas over a relatively short period of time, either by a Popperian negative crucial experiment or a Kuhnian 'Gestalt-Switch'. In the case of the Vitamin Theory the momentum of progress was sustained by the recent converts and not the originators of the theory. In 1909 McCollum had flirted with the flavour enhancing theory, in 1911 he had supported the inorganic salt theory and in 1913 he had been converted to the vitamin theory. Shortly afterwards it was he and not the long standing proponents of the theory who isolated the first putative vitamin. Perhaps there is a moral to be drawn.

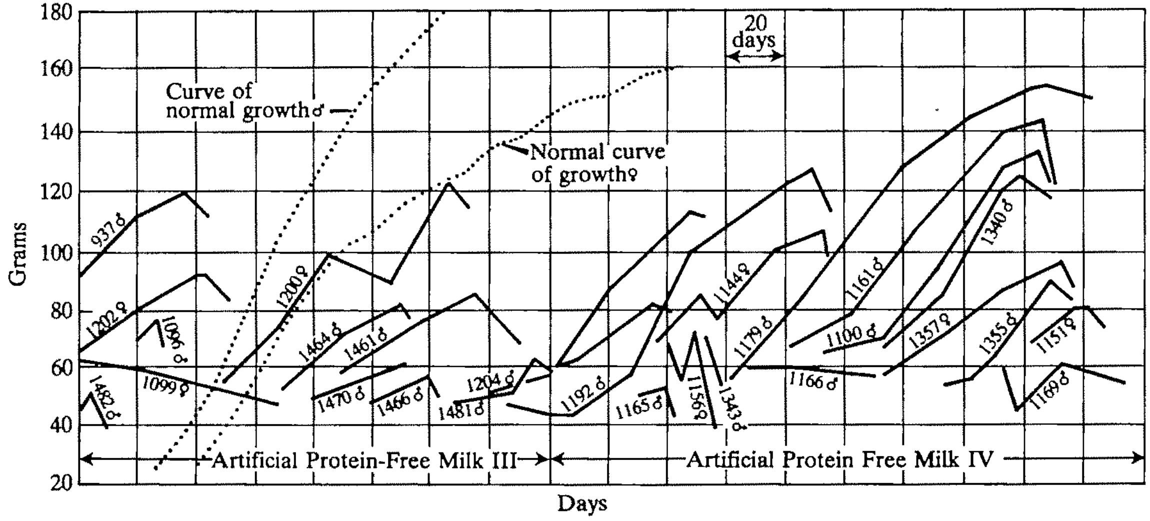


Chart V. Curves of growth of all our rats fed on foods supplying the inorganic constituents in the form of 'artificial protein-free milk' III and IV. Note the very slight growth made on III compared with that made on IV. The essential difference between III and IV is that the latter contained traces of Mn, Al, I and F.

The ordinates represent grams of body weight, as indicated. The divisions of the abscissa represent 20-day periods.

From T. B. Osborne and L. B. Mendel, *Journal of Biological Chemistry*, 1913, 15, 324.

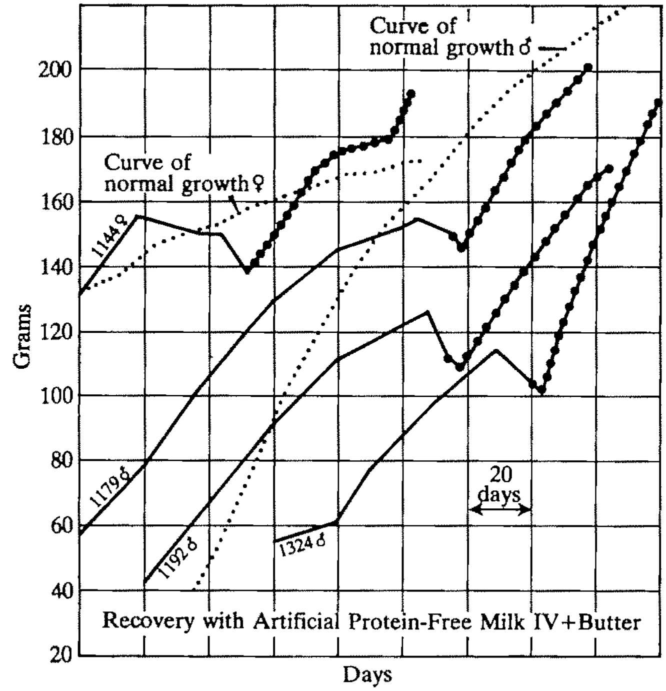


Chart VII. Curves of the body weight of rats which have ceased to grow or have declined on foods containing the 'artificial protein-free milk' IV, and have recovered when part of the lard of the diet was replaced by a corresponding quantity of unsalted *butter* as indicated by the interrupted lines (o-o-o-o-o-o-o). The proteins furnished in the different experiments were as follows: casein, Rats 1144, 1192, 1324; edestin, Rat 1179; lactalbumin, Rat 1144.

The ordinates represent grams of body weight, as indicated. The divisions of the abscissa represent 20-day periods.

From T. B. Osborne and L. B. Mendel, *Journal of Biological Chemistry*, 1913, 15, 326.

REFERENCES

Akeroyd, F. M. [1985]: 'The Relationship between Popper's Philosophy and the Acceptance of the Vitamin Theory', *Journal of Biological Education*, 19, pp. 224–226.

Akeroyd, F. M. [1986]: 'A Challenge to the Followers of Lakatos', British *Journal for the Philosophy of Science*, 37, pp. 359–362.

Hopkins, F. G. [1912]: 'Feeding Experiments Illustrating the Importance of Accessory Factors in Normal Dietaries', *Journal of Physiology*, 44, pp. 425–460.

Lakatos, I. [1970]: 'History of Science and its Rational Reconstructions', PSA 1970, ed. R. C. Buck and R. S. Cohen, D. Reidel, pp. 91–108.

LOTHROP, A. P. [1917]: in *Chemical Abstracts*, **11**, p. 2486.

MACCOLLUM, E. V. and DAVIS, M. [1913]: 'The Necessity of Certain Lipins in the diet during Growth', *Journal of Biological Chemistry*, **15**, pp. 167–175.

OSBORNE, T. B. and MENDEL, L. B. [1911]: 'Feeding Experiments with Isolated Food substances', Publication No. 156, Carnegie Institute of Washington.

OSBORNE, T. B. and MENDEL, L. B. [1913]: 'The Relation of Growth to the Chemical Constituents of the Diet', *Journal of Biological Chemistry*, **15**, pp. 311–326.

OSBORNE, T. B. and MENDEL, L. B. [1917]: 'The Role of Vitamines in the Diet', *Journal of Biological Chemistry*, **31**, pp. 149–163.

Brit. J. Phil. Sci. 39 (1988), 59–65 *Printed in Great Britain*

Chalmers on Method

BARRY GOWER

1. To judge by current methodological writing, the idea of a method characteristic of scientific inquiry is not popular. The imperialism of Aristotelian, of Newtonian, of positivist principles of method, has given way to an eclectic approach in which the task of understanding what various scientists variously do has superseded attempts to pontificate about what all scientists should always do: *tout comprendre, c'est tout pardonner.*

Alan Chalmers, in his recent discussion of Galileo's use of telescopic observations, as well as in the new edition of his textbook *What is this thing called science?* [1982], has endorsed this opposition to scientific method. 'Scientists,' he claims, 'should not be constrained by the rules of the methodologist' (*op. cit.*, p. 135). Fruitful violations of methodological rules cannot, he says, 'be anticipated and legislated for in advance.' New practices will alter methodologies 'and for this reason the notion of a universal, ahistorical account of method that can serve as a standard, not only for the present but also for future knowledge, is an absurdity' (Chalmers [1985], p. 183). Evidently Chalmers is 'against' method even if he is also critical of some of Feyerabend's argumentation.

Some of us, however, have yet to be convinced that sceptical relativism about scientific method has been properly justified, and Chalmers' argument helps to reveal the weakness of some standard moves made by critics.

2. Superficially, at least, the case against method proceeds something like this: some facts about the way science is actually practised are presented as inconsistent with methodological principle. Almost invariably, principle is obliged to give way in favour of practice. So, for example, Galileo's preference for telescopic evidence over the testimony of unaided senses is said to conflict with 'Aristotelian orthodoxy' (*op. cit.*, p. 182), and we are invited to reject the latter as inadequate by accepting Galileo's practice as successful.

The falsificationism, naive or otherwise, inherent in such thinking is obvious. It should be observed, therefore, that inconsistencies between thought, as embodied in general principle, and reality, as exemplified in particular practice, are *made* rather than discovered. With the aid of ingenuity and perhaps audacity, any practice can be reconciled with any principle. Take, for simple illustration, a falsificationist principle itself, namely: Every genuine test of a theory is an attempt to falsify it. Notoriously, those who try to query the principle by reference to what scientists have actually done soon find that any practice is consistent with this principle, and learn that, in fact, falsificationist principles fit the history of science 'like a glove'. For, if we wish to attack it by identifying some test of a theory which is *not* an attempt to falsify it, we cannot do so without deciding that the test in question counts as genuine. Given the provisional nature of such decisions it is clear that a falsificationist could protect his principle by deciding differently about the test.

But the merits or otherwise of conventionalist defences of methodological principle are, in any case, beside the point. It would be wrong to suppose that widespread scepticism about methodology is the product of an implicit empiricism which says that practice matters more than principle in methodology, just as it says that fact matters more than theory in science. In the first place, such an empiricism would be altogether too facile. And secondly, some of the most vociferous critics of methodology are very far indeed from recommending an empiricist attitude in science. They may be among the first to condemn philosophers for ignoring facts about scientific practice, but they are also among the first to congratulate scientists for ignoring the evidence of their senses. There is, surely, something incongruous in the belief that the best methodologists, by paying attention to the historical facts about science, have come to the conclusion that the best scientists ignore facts.[1]

3. As is so often the case with relativistic views about truth, or justification, or rationality, doubts about the coherence of the anti-methodologist's arguments are well-founded. Chalmers, for example, says that he rejects the rule which requires scientists to 'reject theories that are incompatible with generally held facts' on the grounds that it is a rule 'incompatible with those episodes in science that are commonly regarded as constituting its most progressive phases' (Chalmers [1982], p. 135). He does not say how he reconciles his rejection of a rule with his need of it to justify that rejection.

More generally, there is a substantial and insufficiently noticed difficulty for methodological relativism in that by adopting it we rule out any possibility of its being properly supported by evidence. For, it seems clear that the criticism of methodological theory in the light of scientific practice cannot be effective unless that practice is successful. How, though, is this success to be recognised? In particular, can we distinguish between real and merely apparent success? Aristotelian practice must have *seemed* successful to many of Galileo's predecessors, yet no one, I take it, would wish to suggest that Galilean principles of method might have to be abandoned or adjusted in the light of that success. And the reason for this is clear: our own standards and principles are, we believe, closer to those of Galileo than to those of Aristotle, and it is quite natural, therefore, that we should judge the success of Galileo's practice to be, on the whole, real, and the success of Aristotle's practice to be, for the most part, only apparent. But this entails the supposition that our standards are, in turn, closer to some objective truth about these matters. For unless we do suppose some mind-independent standard of correctness here, the distinction between what *seems* right, effective and successful, and what *is* right, effective and successful will collapse. It is hard, therefore, to see how methodological relativism can be properly supported by an appeal to the history of science without covertly assuming an objective conception of methodology. It is true, of course, that this objection shows only that a familiar way of trying to establish methodological relativism is faulty, and not that that doctrine is itself false, incoherent or inconsistent. Nevertheless, an important reason for the doctrine's popularity is its supposed conformity with historical evidence, so if, as claimed, this reason is spurious, then the attractions of the doctrine are liable to fade.

---

[1] For some further discussion of this incongruity see Gower [1984].

4. A further argument against method which may be implicit in appeals to scientific practice is one that mirrors a familiar argument for realism *in* science. The efforts of methodologists to formulate theories of method which would be of any assistance to scientists, or which would help us to explain scientific practice, have been, it is said, conspicuously unsuccessful. We are told, sometimes at wearisome length, that no methodology so far proposed has even come close to enabling us to understand anything other than the coarsest outlines of the authenticating practices used in science. And, so the argument continues, the best explanation for this fact is, quite simply, that there is no such thing as objective, mind-independent, ahistorical truth in methodology. The continued failure of methodologists to achieve even approximate truth would, so to speak, be 'miraculous' if truth in method were achievable. In other words, just as the instrumental efficacy of scientific theories would be difficult to explain unless they can be understood as providing access to observation-transcending facts, so the continued instrumental futility of methodological theory is best explained as a natural consequence of the absence of any practice-transcending facts for such theories to state.

The worth of this argument seems as questionable as the corresponding argument about scientific theories (*cf.* Laudan [1984], p. 218*ff*). Most conspicuously, the kind of argument in use here is not one which can be taken seriously by anyone who adopts its conclusion. For, its conclusion is that there is no such thing as objective truth in methodology, yet the argument invoked manifestly requires this conclusion to be, itself, an objective truth. It may well be a consequence of the non-existence of methodological truth that theories of scientific method should be, in certain contexts and for certain purposes, instrumentally inadequate, but the observed truth of this consequence cannot be a legitimate ground upon which an anti-realist can advance his anti-realism.

5. The sheer fact of disagreement in methodology is another feature which seems to have impressed the sceptics and fostered their relativism. Thus our attention is drawn to the fact that Galileo's judgement concerning the worth of telescopic observations is at variance with an allegedly empiricist insistence that unaided experience is the only proper basis for scientific conclusions. In general, so it is claimed, methodological beliefs can vary dramatically not only between different scientists working at different times or in different subject areas, but even between scientists working at the same time and in the same subject area. The notorious refusal of some of Galileo's colleagues to make use of telescopic observation may be deemed perverse by us, but apposite by most members of their intellectual community. After all, Galileo himself had great difficulty in persuading others that his new device disclosed real celestial phenomena rather than illusory products of the lenses used (Geymonat [1965], pp. 44–45). But, so the sceptical argument continues, such 'disputes' cannot be settled in the absence of an objective standard of rationality. If we declare Galileo's opponents to be irrational we simply expose our allegiance to a conception of rationality which was not shared by those opponents, who would no doubt declare us to lack the kind of rationality which mattered to them. All we can say, it seems, is that conflicting methodological judgements are simply products of distinct systems of methodological beliefs which are coherent within their own terms. We will search in vain for any 'fact of the matter' which could resolve disagreement; Galileo's judgement is not *truer* than an empiricist's for there is nothing for it to be truer to. There are, it is presumed, no methodological facts. So, although there can be change in methodology, there can be no progress for there is no such thing as a true theory about method in science which it is the aim of methodologists to establish.

The issue here is a familiar one from other contexts; it is not easy to tell whether a conflict of beliefs about some topic obliges us to relinquish any idea that the beliefs are answerable to an objective subject matter, or whether, instead, we are faced simply with differences in concepts. Dispute about the validity of, say, a precept involving the use of moral concepts, or the truth of a law using scientific concepts, cannot be assumed to imply disagreement about the truth value of the precept or law; it is at least as likely that it implies, rather, a difference about the scope of the concept involved. In a similar way, it is not at all as clear as Chalmers suggests that a difference of opinion between Galileo and his 'Aristotelian' contemporaries about the worth of telescopic observation implies any disagreement about a methodological precept such as 'it falls to experience to supply the principles of any subject'.[2] A more likely implication is that Galileo differed from some of his colleagues and competitors in what he was prepared to count as evidence, as observation, as fact. After all, the speed and apparent ease with which Galileo extended and developed his concept of evidence is one of the more remarkable features of his achievement.

---

[2] Chalmers [1985] quotes this remark of Aristotle who says that it means for example 'that astronomical experience supplies the principles of astronomical science: for once the phenomena were adequately apprehended, the demonstrations of astronomy were discovered.' (*Prior Analytics*, I, 47a) There is, of course, no suggestion in these comments that the adequate apprehension of astronomical phenomena can only be provided by experience unaided by telescopes.

6. However, even if Chalmers is right in claiming that Galileo's use of the telescope did not just undermine the credibility of a particular scientific theory but also 'constituted a challenge to the standards that [Aristotelian methodology] embodied' (Chalmers [1985], p. 182), it does not follow that methodological realism is to be faulted. In their theories of method, as well as in their practical methodological judgements, people can and do make mistakes. If Aristotle, or anyone else, ever claimed that the verdict of unaided observation was always and everywhere to be preferred to aided observation, then he was wrong. Indeed, Chalmers' tone implies just this by suggesting that Aristotelian principles *should* have given way to Galilean practice. Aristotle's *scientific* views have been challenged yet we do not for that reason abandon realism *in* science, so why should a challenge to his views about scientific method oblige us to discard realism *about* science? Error in methodology, like error in science, can sometimes be explained and can sometimes be corrected by rational means; in neither case is it always and everywhere a matter of the inexplicable replacement of one more or less arbitrary view by another. If Aristotle's mistakes in astronomy, because they can be explained and corrected, do not oblige us to abandon realism in science, why should his explicable and correctable mistakes in methodology threaten realism about science? In short, the evident differences between Aristotle and Galileo about scientific method are no more perturbing to realism than are the equally evident differences between them about, say, the truth of a heliocentric theory of planetary orbits. More generally, if the changes that have taken place in scientific method can be shown to have come about in an explicable and non-arbitrary manner, then relativism is not a good explanation of disagreement in methodology.

7. One further source of relativism in methodology deserves some mention. It is often assumed that methodological rules are relative in the sense that their admissibility or appropriateness depends upon the aims of the scientists subject to them. Since these aims are, or can be, a matter of choice, so also are the methodological rules that embody them. There can, therefore, be no question about the objective validity of these rules. Thus, since Aristotle took the view that scientific inquiry should aim at knowledge of the 'reasoned' fact, we expect an emphasis upon demonstrative reasoning and on formal logical techniques to feature prominently in an Aristotelian theory of method. We expect, too, the concepts of reason, authority, enlightenment, and understanding to be of central importance in such a theory. But, as is now plain to see, there is no obligation to accept this account of what science should achieve.

Euclid's *Elements* and the geometrical science of positional astronomy, despite their intellectual and practical virtues, are in many ways misleading models of what is possible in science. The new and sometimes surprising models that were created in the seventeenth century led eventually to a different conception of science, and to a suitably transformed methodology. Such concepts as hypothesis, conjecture, evidence, probable belief, and ampliative reasoning, which had previously been neglected, shunned, or unknown, began to assume the important role in methodological considerations that they now have. Our requirements are not those of Aristotle and it is a mistake, therefore, to suppose that our methodological standards can in any useful way be compared with his.

But in this case, too, more caution needs to be exercised. We may acknowledge that there have been in the past different *conceptions* of scientific inquiry and its aims. But this is not to say, as required by the relativist, that there are different *concepts* of scientific inquiry and its aims. Though the relativist will deny, he does not refute the realist thought that Aristotelian aims for scientific inquiry have turned out to be too limiting and distorting. Moreover, there needs to be argument for the key assumption that methodological judgements reflect aims in that they can have force and application only for those who decide to pursue those aims. Aims themselves require justification and it is misleading to represent them merely as conditions contributing to the motivational basis for principles of method. Whatever may be the motivation for a scientist's adoption of a methodological rule, it will be appropriate, because of his adoption of it, to ascribe to him the corresponding aims. Accordingly, the explanation for his pursuit of these aims will be the same as the explanation for his adoption of the rule.

It is true, of course, that history reveals diversity in method as well as in science itself, but we should not be so impressed by the diversity that we ignore the connections. It is, after all, we and not Aristotle who label some parts of his work 'scientific'. The belief that scientific truth is created rather than discovered has proved difficult to sustain; we should not suppose that the belief that methodological truth is a variable fiction of our philosophical imagination rather than a matter of how, objectively, things stand will be any easier to defend.

University of Durham

REFERENCES

CHALMERS, A. [1982]: *What is this thing called science?*, 2nd ed., Open University Press.

CHALMERS, A. [1985]: 'Galileo's Telescopic Observations of Venus and Mars', *British Journal for the Philosophy of Science*, 36, pp. 175–183.

GEYMONAT, L. [1965]: *Galileo Galilei: A Biography and Inquiry into his Philosophy of Science*, McGraw-Hill.

GOWER, B. [1985]: 'Method in methodoloy', *Methodology and Science*, **18**, pp. 30–47.

LAUDAN, L. [1984]: 'A Confutation of Convergent Realism', in J. Leplin (ed.), *Scientific Realism*, University of California Press.

Brit. J. Phil. Sci. 39 (1988), 67–79 *Printed in Great Britain*

Real Dispositions in the Physical World

IAN J. THOMPSON

ABSTRACT

The role of dispositions in the physical world is considered. It is shown that not only can classical physics be reasonably construed as the discovery of real dispositions, but also quantum physics. This approach moreover allows a realistic understanding of quantum processes.

1 *Introduction*
2 *General Analysis of Dispositions*
3 *Are Dispositions Real?*
4 *Dispositions in Classical Physics*
5 *Dispositions in Quantum Physics*
6 *Conclusions*

I INTRODUCTION

It is a common belief that modern science does away with those obscure notions of 'disposition' and 'potentiality', in favour of an analysis of the component structure of the things concerned, and their functional relationships. Science, it is often said, cannot long accept an explanation of an object breaking in terms of just its 'fragility', or of a plant seeking light in terms of just its 'phototropism'. Dispositions, so popular opinion has it, are regarded by the scientist as merely a sign that he has to work harder, to find the underlying structural forms and their relations. Talk of 'dispositions', 'powers', and 'capacities' is somehow not seen as sufficiently *definite* for a hard-nosed scientific explanation. There seems to be something intrinsically unsatisfactory and vague about a property that *may or may not* operate, and in particular it seems uncertain how to describe them rigorously and mathematically.

Suppose we do follow the above proposal, and analyse observed dispositions in terms of constituents. Presumably the parts are to have the *abilility to interact*. But this means at the microscopic level of explanation we again have to accept some kinds of dispositional properties: of the parts this time. This is because 'to be able' signals a dispositional property. Thus I will argue (somewhat controversially) that in fact dispositional properties, though perhaps explained, are never explained away. However much we may dislike them, they are never found to be illusory, and cannot be completely replaced by talk of functional relationships, differential equations, and probability calculus.

At the microscopic level we *might* hope that the constituents have *many* definite properties (e.g. mass, shape, position, velocity, energy etc.), and only a *few* of those peculiar dispositional properties (e.g. perfect elasticity, gravitational attraction, and electric charge). In that way there might be a minimum number of these inexplicable 'occult powers'. Such would be the case if Newtonian physics were true, as seen in section 4.

Quantum phenomena show, however, that this hope is not satisfied. Section 5 will describe how in quantum physics there are *more* kinds of dispositions than in Newtonian physics. For the properties of position and velocity, previously thought quite definite, now behave like dispositional properties that may or may not have definite values. I will be describing the 'propensity interpretation' of quantum physics, and it will be seen that in the quantum world there are very few *non*-dispositional properties, i.e. very few properties that always have perfectly definite values. In particular, it will turn out that there is no such thing as a definite corpuscle: following Maxwell [1985], this concept is replaced by that of a 'packet of propensity' or 'propensiton'. Indeed, relativistic quantum field theory goes on to describe how even the *existence* of such an ultimate constituent is itself a probabilistic notion.

2 GENERAL ANALYSIS OF DISPOSITIONS

I will be using the terms 'power', 'potential', 'capability', 'capacity', 'propensity' and 'cause' as examples within the category of 'dispositional properties of objects'. The ascription of properties in this category is typically, adapting a definition of Harre and Madden [1975, 1970], of the form

> 'Object S has the dispositional power P to do action A'
> if and only if
> 'if S is in some circumstance C, C depending on P and A, then there will be a non-zero likelihood of S doing A, in virtue of the constitution of S'.

Here, the 'circumstance C' is usually defined by multiple spatial relations to other objects, and the 'action A' can either be a change in S itself or an interaction with other objects. Harre and Madden [1975] designed the phrase 'in virtue of the constitution of S' to exclude 'changes' to certain properties of S that are changes in purely external relations that may come about completely independently of whatever S is actually like. Thus, for example, no disposition of Socrates is necessary to explain his becoming smaller than Theaetetus, if it is the latter who is growing.

Note that we are not with these ascriptions assuming a permanent 'nature' of the object from which all its powers can be always deduced, only that at any given time there is 'something about the object', a 'real internal constitution' (to use Locke's phrase) that explains all the dispositions it in fact has at that time.

A distinction[1] is thus made between the 'Principal Cause' (that disposition which operates), and the 'Instrumental Cause' (that circumstance by means of which dispositions operate). Principal causes operate according to instrumental causes. Both are necessary for any action, for example, when a stone is let fall: the principal cause is the earth's gravitational attraction, and the instrumental cause is our action of letting go. Its hitting the ground is thus caused by our letting go, but only as an instrumental cause. Many common uses of 'cause' refer to instrumental causes rather than principal causes, as it is only in the instrumental sense that *events* can be said to be causes.

The next point to note is that dispositional properties can only be explained or reduced to other dispositions, *not* to entirely static or structural properties. That is, dispositions have a 'categorical irreducibility', as it is impossible to explain them away in terms of other categories such as space, time, form, process, material, property etc. For suppose that the exact shape and size of an object were known, the shapes and sizes of all its constituents, along with a list of these facts at every time. We would still know nothing about *how* or *why* the object would change with time or on interactions. Still less could we predict how it would respond to a new experimental test. In fact, if it and its parts had no dispositional properties, as Hume would argue, then we have his conclusion that any actions or changes (apart perhaps from uniform motion) would be entirely inexplicable: there would be nothing about the object that could lead to these changes rather than to any others. This categorical irreducibility was seen clearly by Aristotle and Leibniz, as discussed by Leclerc [1972], and has been explained at some length recently by Weissmann [1965], Mellor [1974], Harre & Madden [1970, 1975], and Emmet [1984], among others (see Tuomela [1978]). According to Shoemaker [1979], the continued identity of objects also depends on their causal properties.

An opposing view has recently been argued by Prior [1985], who claims that a disposition must ultimately be reducible to a non-dispositional *basis*. She claims for example that, since the laws of physics are contingent, we can imagine possible worlds where, say, a body with inertial mass would not experience finite acceleration under any applied force (i.e. it would not satisfy the usual subjunctive for inertial mass). This would appear to indicate that inertial mass could not be identified with the dispositional property of being such that the subjunctive is always true, but must instead have (or be) a non-dispositional basis. In this latter account, inertial mass is whichever property is responsible (in just this world, *or* in all worlds) for being such that the subjunctive conditional holds.

---

[1] However, this distinction does *not* simply match Aristotle's distinction of material, efficient, formal and final causes.

The difficulty with this argument is that I cannot see how, in *any* possible world, a purely non-dispositional basis can ever be *responsible* for a dispositional property, in the sense of *implying* the disposition. The (very weak) sense of 'responsibility' she has invoked is practically 'by stipulation', this being apparently the nature of physical contingency. That is, physical law 'just says' that a certain static property called inertial mass is (somehow) 'responsible' for a certain subjunctive condition. This, however, appears to make physical law essentially arbitrary, and makes realistic interpretations difficult. I believe that a more satisfactory account of physical contingency is given by Maxwell's [1968, 1985] 'conjectural essentialism' (see also Harre & Madden [1975]). In this account, dispositional properties such as inertial mass necessarily have their associated conditional property, but it is a completely contingent and empirical question whether any given body (or any body at all) has that kind of inertial mass. Maxwell's account (by having bases that are intrinsically dispositional themselves) gives fundamental bases a much stronger sense of 'responsibility' for observed dispositions. It is precisely because there are such weak connections between purely static and dispositional properties that Prior's purely non-dispositional bases are unsatisfactory, as dispositions cannot be properly explained by static properties. Inertial mass, for example, must be either implied by or identical with a subset of the basic and fundamental *dispositional* properties.

3 ARE DISPOSITIONS REAL?

There seems no way to avoid *some* kind of irreducibly dispositional properties of physical objects unless we find force in one or more of the following objections: (a) that one should deny the subjunctive in the power ascription as anything more than hypothetical, (b) science should be content with finding only the regularity of effects, and not try to discover causes and determining constitutions, or (c) the world should be considered as a 'Zeno universe' that has only successive states and no proper changes.

Taking approach (a), Ryle [1949] denies that dispositional ascriptions 'assert extra matters of fact' and claims that they are *only* 'inference-tickets, which licence us to predict, retrodict, etc.'. That is, he would omit the 'in virtue of the constitution of S' phrase in section 2. Since then, there is no property of S which makes the ascription true, that truth cannot be explained by properties of S. Thus Ryle (quite explicitly) denies that one should even look for either causal or mechanistic explanations of the dispositions. Even, presumably, in cases in physics and chemistry where there are quite obviously explanations in terms of constituents and their propensities to attract and repel each other. His restriction on looking for explanations in terms of internal dynamics is largely disregarded in scientific practice.

Concerning injunction (b), I accept that all observations are of effects rather than of causes, but that does not mean we cannot conceive of causes and of the way they might lead to effects. It seems to me quite legitimate for causes to be postulated in a Popperian fashion, and the consequences deduced for the production of effects. The reverse *in*duction is quite different: obviously we cannot deduce precise causes from observable effects (as Hume has long pointed out), but that does not mean that there no causes. It should not be necessary to accept Hume's [1739] conclusion that 'the distinction, which we often make between a power and the exercise of it, is without foundation'.

To consider (c) the world as a 'block universe' or 'Zeno universe', as recently pointed out by Emmet [1984], is to see only different states of affairs at successive times, and not to see the changes that lead to these differences. Since the rise of Einsteinian relativity, it has become popular to see all of time and space in one 'block continuum' of four dimensions, and to see change as only the difference between successive 'time slices' of this continuum. In this world there is only what actually happens, and as in (a) above, what 'might have happened' is purely hypothetical. The only sense of 'might happen' that can be invoked is to imagine an entirely new possible world, e.g. one with different initial conditions or different physical laws. This world view thus does not base power ascriptions on any real features of this particular universe.

If we reject the 'block universe', we are also rejecting the account of time in which the future is 'already formed' and perfectly definite in advance of its happening, if indeed on this account anything happens at all. (In extreme versions of this theory, time and real change are both completely illusory.) I admit that these accounts have an internal consistency that makes them difficult to refute, but despite their being advocated by many philosophers and physicists, I do not believe that they should be the only coherent metaphysical systems on offer. I believe that alternatives can be devised that not only have more explanatory power (e.g. for quantum physics, see section 5), but also more practical use (see below, end of this section). Once these alternatives have been formulated, we should be able to decide in which way our world is more adequately described.

Rejecting the objection (c), however, does *not* mean rejecting all talk of the spacetime continuum. One alternative approach (following Maxwell [1968, 1985 sections 5 & 6]) says that it is facts about objects, rather than objects themselves, that are (or can construed to be) spread out in time. Spatial relations are between objects, but temporal relations are between facts about objects. Another view follows Whitehead in regarding the spacetime continuum as giving the ordering of possible events (even before they are actual).

Reitdijk [1966, 1973] and Maxwell [1985 sections 1–4] have claimed that special relativity is *in*compatible with the real actualisation of dispositions, as a global account cannot be given of which dispositions are actual and which are not, at any single (metric) time. But that there is no 'global metric time' does *not* mean that we cannot conceive of a 'process time' that counts actualising changes where ever they may be occurring.[2]

It does in any case seem very odd to deny that objects have dispositional properties that relate what might happen as well as what actually happens. To deny causes apart from their manifestations, Mellor [1974 p173] shows, leds to some bizarre consequences, for then dispositions could not be ascribed when their displays are impossible. Taking precautions to avoid the conditions in which nuclear fuel would explode, to use his example, should not mean that the fuel was not explosive. 'It is ridiculous to say that their success robs the fuel of its explosive disposition and thus the precautions of their point.'

Furthermore, a theory that only predicted what actually happens, and not what might happen, would be useless as an engineering planning tool. For it would not be able to predict the consequences of a plan that we might consider employing, but in the end did not actually use. Such a theory could not tell us, for example, what would have happened if one of the *unsuccessful* channel-tunnel designs had been chosen. These questions are of great practical (and political) importance, and real physical theories are of course most useful here.

4 DISPOSITIONS IN CLASSICAL PHYSICS

In this section, we show first that dispositions (of kinds to be determined) have an essential role in physics, and secondly that mathematical physics is an attempt to relate forms and dispositions in a regular manner. Physics need not be concerned with the ultimate nature of dispositions, but only with knowing that they *do* exist, and then with investigating their properties, locations, interactions, effects, changes with time, etc. in as much detail as possible. We will see that dispositions are important in almost all kinds of Newtonian physics.

Dispositions first appear in physics as the macroscopic features of observable objects that we wish to explain. Dispositional properties are largely those which cannot be explained purely by the location and shape of these objects, but require *causal* kinds of ascriptions and analyses, as explained in section 2, in terms of causal powers. Thus, if we wanted, one way to deny causal powers would be to deny the reality of *all* dispositional properties, as discussed in the previous section, by denying the reality of the subjunctive in the power ascription. On this account, what *might* happen to objects is completely irrelevant to what *really* happens. If a glass is never going to be hit, for example, it is purely hypothetical whether it is ever able to break. In physics, as in politics, hypothetical questions need not be answered. They need not be, but they usually are: we would hardly be satisfied with a physical theory, however true, that only holds for laboratory test cases. And if they are answered, on the basis of which real properties are the answers correct, if not dispositional properties?

---

[2] For the concept of 'process time' as a 'universal passage of nature' see for example Capek [1961], pp. 205 *et seq.* What has to be done now is to understand how process time is an 'extensive becoming' (Capek [1961], p. 220). This 'extension' is necessary to explain the 'reduction of the wavepacket' in relativistic space, and is the subject of further investigation.

Another way in which the reality of causes could perhaps be denied is to say that physics is *only* the discovery of laws that relate events, not the explanation of the properties of things that lead to these events: that is, that physics is (or should be) only concerned with effects, not with causes. It is agreed that all observations are effects of interactions, but it does seem an unnecessarily severe restriction not to permit physicists to speculate on the causal properties of what they are examining, nor to permit them to postulate, for example, potential energy apart from kinetic energy. Without potential energy, as in a coiled spring, we could not even have the conservation of energy.

A mathematical physicist might argue that what is objective and essential about a physical theory is a certain differential equation and its family of solutions in phase space. I agree that it would be a significant scientific achievement to have devised such an equation that correctly predicts events and movements etc. But if we are not satisfied with a purely instrumentalist account of scientific theories, we can still ask 'what is it about the physical world that makes this theory a correct description?'. Why does the differential equation correctly predict the world? Perhaps because it describes correctly the time evolution of dispositions such as forces, potentials, or quantum wave functions, in which case it can be construed as a good description of these real dispositions. Perhaps, however, it cannot be simply interpreted in this way, yet still gives good predictions of what actually happens. I would then ask whether the equation can also describe what particular objects or particles *might do* if put in new situations. If it cannot, then it clearly falls short of the tasks physicists set themselves, and it would be useless for planning in engineering. If it does answer hypothetical questions, I ask: which properties of the objects involved determine the parameters of the equation? How, for example, is the equation set up with the correct strengths of couplings to the gravitational, electromagnetic and/or nuclear fields? I claim that it is precisely these dispositional properties of the objects that are needed to set up hypothetical test cases, with purely structural properties being inadequate. That is, it would appear that any mathematical description can answer all hypothetical questions correctly *only* if it has embedded within it (explicitly or implicitly) a description of causes and their operations. The equation, for example, would have different structures if vector influences combined in different ways (e.g. by vector sum, or by largest effect dominating, or by random selection of effects from different sources).

A fourth kind of physical theory that appears to avoid real dispositions is

Einstein's general theory of relativity, and its modern variant, geometrodynamics. Both of these theories are usually interpreted in terms of a 'block universe' with time as if spatially extended, as they reduce all dynamics to the geometric properties of spacetime. As however both theories can and have been used to answer hypothetical questions (e.g. what would happen to the solar system should the sun suddenly become deformed), they can both be reformulated as sets of rules for deriving the state of the universe (along some spacelike hypersurface) at some time t′, given its state at some earlier time t<t′. In a realistic interpretation of this time-dependent formulation, it turns out that dispositions (or something very similar) reappear. For now general relativity describes how the mass-energy tensor causes spacetime curvature, and the curvature itself in turn describes how objects *would* move in spacetime *if* they were present. That is, matter can be regarded as influencing the dispositions of objects to move in straight or curved paths.

Geometrodynamics (assuming the theory can be worked out in adequate detail) has a similar 'thick sandwich' time-dependent interpretation, but is different from general relativity in that now physical objects are not *in* spacetime (as we have always imagined), but simply *are* regions of spacetime with certain patterns of curvature. In the time evolution from t to t′, these patterns interact in a non-linear fashion, and attract and repel each other as do physical objects. But this means that implicit in the non-linear field equations are rules for determining how a given pattern of curvature would interact in various circumstances, on the basis of its nature as a pattern. That is, these patterns do have dispositional properties. I agree that in this theory the nature of objects would be a set of structural properties of spacetime curvature, and not a set of dispositional properties. But since these structural patterns only have significance in conjunction with the field equations, and because this conjunction results in true dispositions, the theory is not incompatible with reality of dispositions. It is just that, if geometrodynamics were correct, the dispositions would be properties of spacetime itself, not of physical objects in spacetime (as there are no such things).

A significant part of ordinary physics is the explanation of macroscopic dispositional properties in terms of the dispositional properties of the components and the configuration of these components in the whole. Thus the elasticity of a solid, for example, is explained in terms of the attractions between the electrons and their neighbouring atoms. Note that it is *not* enough to say that the elasticity can be explained simply in terms of the 'electronic structure', as purely structural properties cannot explain dispositional features without assuming some dispositions (such as charge, mass etc.) inherent in the electrons themselves.

There have been very few attempts in physics to deny that the constituent parts *do* have causal properties—i.e. that electrons do not really have electric charge, mass, and spin. Even the proposed quarks have these dispositional properties, along with 'colour charge', 'strangeness', 'charm', etc. It is an empirical question *which* causal properties are basic, but physicists nearly always rely on some causal properties being fundamental. The form of the postulated basic causal powers (nuclear and electric fields, etc.) may be constrained by mathematical laws of symmetry, so that for example there are only discrete values for charge, spin etc., but the *existence* of these causal powers is something that *must* be assumed in order to provide a basis for physical accounts of the dispositions evident in nuclear, chemical, and biochemical systems.

The task of physics must therefore be to relate causal properties to what is known about the actual forms of the objects under investigation. Mathematical physics describes the numerical features of natural objects, and simply assumes that there are dispositional properties that exist according to the described forms. For example, the attribution of electric charge is purely formal until it is assumed either that there is a real dispositional property (e.g. a force) with corresponding features, or that there is a corresponding coupling to a potential-energy field. ('Forces' and 'potentials' are equivalent descriptions of the same disposition, as forces are spatial gradients of potential energy field.)

In Newtonian physics the mechanical corpuscules had the dispositional properties of impenetrability, durability, and perfect elasticity according to their spatial shape. These causal properties cannot, however, be logically derived from the shape alone, as Descartes was forced to acknowledge. Newton also realised that other causal powers must be attributed to the atoms, in order for example to explain gravity and the tensile strength of materials. Gravitational attraction could be made according to mass and distance, but the short-range attractions could not be attributed according to any form known at that time. The fact that the attribution of gravitational powers was in strict accordance to some known form may have contributed to the impression that gravity was satisfactorily explained, but this attribution does not remove the need for a dispositional category. For Newton's law of gravitation does not say what always *does* happen, but only what *would* happen in suitable circumstances (e.g. no interference from outside influences: in Newton's case, no nuclear or electromagnetic forces). This fact was one of Nancy Cartwright's [1983] objections to the reality of physical laws, but is understandable when laws, even Newton's law, are regarded as laws of causes, not simply as laws of effects.

5 DISPOSITIONS IN QUANTUM PHYSICS

Position and velocity were also spatial properties in Newtonian physics. That is, they were definite and actual features of the elementary corpuscules, and could be called 'definite' rather than 'dispositional'. This belief changed with the advent of quantum mechanics, which showed that position and velocity are not continuously definite, but only have specific values in suitable situations such as measurement interactions of certain kinds. That is, position and velocity values are now not 'primary qualities' which exist continuously in all situations, but are more akin to 'secondary qualities' which are only specific in some suitable (i.e. not all) situations. In fact, the circumstances for positions being definite are *incompatible* with the circumstances for velocity being definite: this is part of the content of Heisenberg's 'Uncertainty Principle'. If we take an ontological (rather than epistemological) view of what this principle means, then position and velocity values can only be real if they are features of actions by dispositions, and not definite 'primary' forms themselves.

This shows that position and velocity should now be related *not* to spatial properties or actual shapes, but to *propensities* (Popper [1967], Whiteman [1971] and Maxwell [1982, 1985 appendix]). Quantum mechanics uses the 'propensity' type of disposition, as this type displays its effects probabilistically: see Popper [1959]. If we then ask what must the world be like in order that quantum mechanics describes it correctly, we arrive at the existence of real propensities. It is sometimes thought that talk of propensities can be reduced to talk of conditional probabilities etc., but as Humphreys [1985] shows, this cannot be carried through as propensities have for example a time asymmetry not shared by conditional probabilities.

If we look at what definite forms natural things *do* have in quantum physics, these, I believe,[3] are

1) the set of 'quantum numbers' (for charge, spin etc.) of the quantum systems, and
2) the set of definite past interactions (by measurement or other events)[4] which determine the current quantum state.

Thus in quantum theory, 'what is actual' refers not to spatial shapes, but to combinations of numbers and past events. Within this scheme, quantum objects have propensities according to these definite numbers and events. It is from those corresponding propensities (not the events etc. themselves) that both the subsequent spatial shapes and later behaviour of the objects can be derived. Note that quantum objects do *not* behave as if they had a *definite* spatial shape, because the spatial distribution of propensities varies with time. They are like 'wave packets' of variable extent. Their extension can even have temporary gaps in it, as in the two-slit diffraction experiment.

Quantum mechanics itself would describe the subsequent development of a dispositional state in terms of an evolving Schrodinger wave function, or in terms of an Heisenberg transition operator from the previous definite interactions to the range of possible interactions in the future. These quantum-mechanical procedures are thus an attempt to describe the time evolution of the propensities of natural objects.

---

[3] I am not yet allowing for theories of 'spontaneous symmetry breaking' which, if wholly successful, would account for quantum numbers in terms of actual events in the early stages of the 'big bang'.

[4] There have been several proposals for how events (other than measurement events) can give definite selections, or 'reduction of the wave packet'. See e.g. Maxwell [1982, 1985], or D. Bedford and D. Wang [1975].

Because of this variation of shape with time, and because the 'actual particle' *never* appears (neither in propagations nor in interactions), it is now best to regard quantum objects as essentially just *distributions of propensity over spacetime*. Such objects Maxwell [1985] has called 'propensitons'. The propensities involved are not arbitary dispositions, but are propensities *for* interactions: for localised actual events.[5] This means that *during interactions* the 'propensitons' behave almost as if they were particles, but not at intermediate times. Between interactions, a propensiton does have a spatial extent, but in a probabilistic and variable, not definite, manner.

Of course, many physicists have continued using the term 'particle' for these propensity distributions. The meaning of the word 'particle' has thus changed. Kaempffer [1965], for example, after pointing out the 'erosion of naive pictures of particles', goes on to suggest that the word 'particle' stand for a quantum mechanical state characterised by a set of quantum numbers, which is associated, in principle, with an identifiable event such as a momentum transfer in a collision. Note how this concept is similar to that of a distribution of propensity associated, in principle, with an actual event such as a collision.

In quantum field theory (a more complete form of quantum physics), even the *existence* of objects is a dispositional property that may or may not be manifested, as, for example, pairs of particles and anti-particles may or may not be formed. Quantum field theory furthermore replaces the 'potential energy fields' of ordinary quantum mechanics with the 'virtual exchange of particles'. Both of these are dispositional: virtual exchange is the propensity of generating short-lived particles that may or may not interact with others in their vicinity, and that can also form what physicists often call a 'cloud of possibilities' if they do not interact.

6 CONCLUSIONS

By means of the arguments presented above, I hope to have shown that the notion of a 'real disposition' is an essential part of both our theoretical and practical understanding of the physical world. Not only this, but that the notion is likely to be fundamental to a realistic and non-paradoxical account of quantum physics. For at least these reasons it is thus important to resist certain interpretations of physics and of the physical world that render dispositions impossible.

---

[5] A more complete account of the concepts of 'actuality' and 'potentiality' involved here is in preparation.

In particular, it is necessary to reject the 'block universe' account of time which sees the future as 'already formed', only waiting for us to 'come across it'. That in the future some event is going to happen does *not* imply that there 'now' exists some event that in the future is going to happen. Until events happen, they need not even exist.

Equally importantly, it is necessary to reject the account of physical processes as the mere succession of events, without their being in any sense caused or produced or generated by appropriate previous event(s). This Humean account of process removes by definition the possibility of any rationally coherent account of *how* events are caused.

I recognise that in rejecting the above two accounts, I am also rejecting the simplest interpretation of Einsteinian relativity theory: the interpretation in terms of a single spacetime of fixed events. Since Minkowski, relativity theory is often taken as *implying* this view of spacetime, but I intend to show in subsequent papers that there are more plausible interpretations.

Often allied to the Humean account, and hence also to be rejected, is the conception of physics as finding only the *laws or regularities* that describe the succession of events, or of states of the world. That is, physics is seen as finding functional relationships between events, and *not* the causes of events. Physicists may say that their laws 'govern' the succession of events, but, once the 'block universe' account of time is rejected, it is quite mysterious *how* this can occur. (They may also say that they are 'iron laws', but there is no iron hand of the law to enforce them.)

Such difficulties are avoided if physics is seen as discovering the dispositional properties of objects, so that the above event laws are *consequences* of the way things are, not *fiats* to be enforced. On this construal physics, even Newtonian physics, can realistically describe what 'would happen' in different possible worlds, not just the (simpler) knowledge of what 'does happen' in the present world: it can predict for example what *would* happen in various experiments, even if those experiments are never in fact performed. It tries to do this on the basis of describing the real dispositions of physical objects, and how these dispositions lead to various effects in different circumstances.

University of Bristol

REFERENCES

BEDFORD, D. and WANG, D. [1975]: 'Toward an Objective Interpretation of Quantum Mechanics', *Il Nuovo Cimento*, 26B, pp 313–325.
CAPEK, M. [1961]: *The Philosophical Impact of Contemporary Physics*, Van Nostrand.
CARTWRIGHT, N. [1983]: *How the Laws of Physics Lie*, Clarendon Press.
EMMET, D. [1984]: *The Effectiveness of Causes*, Macmillan.
HARRE, R. [1970]: 'Powers', *British Journal for the Philosophy of Science*, 21, pp. 81–101.

HARRE, R. and MADDEN, E. H. [1975]: *Causal Powers: A Theory of Natural Necessity*, Blackwell, p. 86.

HUME, D. [1739]: *Treatise of Human Nature*, Book I, section XIV, Collins, Fontana Edition, 1962.

HUMPHREYS, P. [1985]: 'Why Propensities Cannot Be Probabilities', *Philosophical Review*, **94**, pp. 557–570.

KAEMPFFER, F. A. [1965]: *Concepts in Quantum Mechanics*, Academic Press.

LECLERC, I. [1972]: *The Nature of Physical Existence*, Allen & Unwin, Muirhead Library of Philosophy.

MAXWELL, N. [1968]: 'Can There be Necessary Connections between Successive Events?', *British Journal for the Philosophy of Science*, **19**, pp. 1–25.

MAXWELL, N. [1982]: 'Instead of Particles and Fields: A Micro Realistic Quantum "Smearon" Theory', *Foundations of Physics*, **12**, pp. 607–631.

MAXWELL, N. [1985]: 'Are Probabilism and Special Relativity Compatible?', *Philosophy of Science*, **52**, pp. 23–43.

MELLOR, D. H. [1974]: 'In Defence of Dispositions', *Philosophical Review*, **83**, pp. 157–181.

POPPER, K. R. [1959]: 'The Propensity Theory of Probability', *British Journal for the Philosophy of Science*, **10**, pp. 25–42.

POPPER, K. R. [1967]: 'Quantum Mechanics without "The Observer" ', in *Quantum Mechanics and Reality*, M. Bunge (ed.), Springer, 1967.

PRIOR, E. [1985]: *Dispositions*, Aberdeen University Press.

REITDIJK, C. W. [1966]: 'A Rigorous Proof of Determinism Derived from the Special Theory of Relativity', *Philosophy of Science*, **33**, pp. 341–344.

REITDIJK, C. W. [1973]: *On the explanation of matter wave interference*, Van Gorcum, chapter 1.

RYLE, G. [1949]: *The Concept of Mind*, Hutchinson, ch. 5.

SHOEMAKER, S. [1979]: 'Identity, Properties and Causality' in *Midwest Studies in Philosophy*, Vol. IV, University of Minnesota, pp. 321–342.

TUOMELA, R. (ed.) [1978]: *Dispositions*, D. Reidel.

WEISSMAN, D. [1965]: *Dispositional Properties*, Southern Illinois University Press, ch. 4.

WHITEMAN, J. H. M. [1971]: 'The Phenomenology of Observation and Explanation in Quantum Theory', in *Quantum Theory and Beyond*, Ted Bastin (ed.), Cambridge University Press.

Brit. J. Phil. Sci. 39 (1988), 81–109 *Printed in Great Britain*

Philosophical Objections to the Kinetic Theory*

JOHN NYHOF

ABSTRACT

Towards the end of the 19th century there were those who wished to see the kinetic theory abandoned. This paper attempts to show that this reaction was primarily due to philosophical objections rather than the result of scientific difficulties encountered by the kinetic theory. First the relevant philosophical background is examined as well as the relation between the kinetic theory and thermodynamics. Next the scientific difficulty known as the specific heats ratio anomaly is discussed and finally Boltzmann's philosophy of science is examined.

1 *Introduction*
2 *Instrumentalism*
3 *Positivism*
4 *The Kinetic Theory and Thermodynamics*
5 *The Specific Heat Anomaly*
6 *Boltzmann and Posittivism*

I INTRODUCTION

The kinetic theory and thermodynamics are important theories developed in the 19th century to deal with macroscopic thermal phenomena. In 1878 Maxwell defined thermodynamics

> . . . as the investigation of the dynamical and thermal properties of bodies, deduced entirely from what are called the first and second laws of thermodynamics without any hypotheses as to the molecular constitution of the bodies ([1878], p. 258).

The kinetic theory on the contrary, did put forward hypotheses as to the constitution of matter. By means of classical mechanics it attempted to deduce the macroscopic thermal phenomena from the motions and collisions of minute particles or atoms.

In the 19th century thermodynamics was universally accepted as a valuable theory. Not so the kinetic theory, attitudes to which ran the whole gamut. As we shall see, Boltzmann was in love with it, Duhem thought it a parasite, and many others were ambivalent about it.

---

* I am grateful to Alan Musgrave for helpful comments.

Until fairly recently historians and philosophers of science seem to have generally accepted that those 19th century scientists who fought to have the kinetic theory abandoned were expressing their positivistic philosophic views. For instance, Brush writes that:

> In retrospect it seems clear that the criticisms of the kinetic theory in this period were motivated not primarily by technical problems, such as specific heats of polyatomic molecules but, rather by a general philosophical reaction against mechanistic or 'materialistic' science and a preference for empirical or phenomenological theories, as opposed to atomic models ([1976], p. 245).

Recently, however a different view has been proposed. It has been claimed that most of the opposition to the kinetic theory must be attributed to the *scientific* difficulties it encountered, particularly when these were contrasted with the successes of thermodynamics. According to one view (Clark [1976]), criticism of the kinetic theory did not find its origin in a positivistic philosophy. Rather it was the other way round: the degeneration of the kinetic theory combined with the progress of thermodynamics gave rise to scientific positivism. Another view (Gardner [1979]) has it that there was a gradual transition from an instrumentalist to a realistic acceptance of the atomic theory as its predictive power increased.

So there is considerable difference of opinion as to the relative influence of philosophical views and scientific difficulties respectively on the development of the kinetic theory and its relation to thermodynamics. Our aim is to show that the older view—the wish to see the kinetic theory abandoned was due to the philosophical objections rather than to scientific difficulties—is the more correct one.

We will argue that those who wished to see the kinetic theory abandoned as a descriptive and explanatory theory did so because they were anti-mechanists. Furthermore many anti-mechanists derived their anti-mechanism from a positivistic philosophy or a certain kind of instrumentalism. It is one thing to raise genuine scientific objections to the kinetic theory because its mechanical explanations appear to have run into serious scientific difficulties; it is quite another to object to the kinetic theory because it makes use of mechanical explanations which are deemed to be philosophically objectionable.

In this paper we will first look at instrumentalism and then at positivism. In these two sections there will be little need to deal with specific difficulties because it will become apparent that these philosophies are more concerned with the general nature of science. It is the descriptive and explanatory nature of scientific theories like the kinetic theory which is thought to be philosophically unacceptable.

Only then, having established part of the relevant philosophical background, we will examine the relation between the kinetic theory and thermodynamics. In this section we will briefly show why phenomenological thermodynamics was acceptable to the positivists, while statistical thermodynamics was not, and how, because of lack of observational evidence for the statistical nature of the second law of thermodynamics, the result was a stalemate between realists and positivists.

Next we examine the specific heats ratio anomaly in some detail because it was the single most protracted and serious scientific difficulty encountered by the kinetic theory. We mention the work of Wüllner and Strecker, which does not appear to be well known, because it helped to convince Boltzmann that certain molecules under particular conditions, really do behave as rigid bodies. However, there does not seem to be any evidence to suggest that this anomaly contributed to the rise of positivism. But even if specific difficulties of a scientific nature were not the *cause* of the rise of scientific positivism, there remains the possibility that certain realists, because of such difficulties, eventually abandoned realism in favour of positivism once the latter had established itself as a philosophy of science. Claims along these lines have been made about Boltzmann. It is generally agreed that Boltzmann for a considerable period of his career was the champion of realism and in particular of the realistic interpretation of the kinetic theory. However, it has been claimed that scientific difficulties led Boltzmann to abandon not only atomism but, far worse, realism; and so he is supposed to have become a positivist for the last decade or so of his life. As Boltzmann is such a central figure we devote the entire final section to these matters.

2 INSTRUMENTALISM

A certain type of instrumentalism holds that scientific theories are not explanations. The originator of this type of instrumentalism is Pierre Duhem (1861–1916), who gave the following definition of a physical theory:

> A physical theory is not an explanation. It is a system of mathematical propositions, which aim to represent as simply, as completely, and as exactly as possible a set of experimental laws ([1906], p. 19).

To the question what it is to explain Duhem replies: 'To explain . . . is to strip reality of appearances covering it like a veil, in order to see the bare reality itself (ibid., p. 7). He continues:

> For such an inquiry to make sense or to be at all possible, we must first of all regard as certain the following affirmation: under the sensible appearances, which are revealed in our perceptions there is a reality distinct from these appearances (ibid., p. 9).

Duhem believes that even if there is a hidden reality it will remain largely inaccessible to science and so theories can at best only provide doubtful explanations.

> Most often we find that physical theory cannot attain that degree of perfection; it cannot offer itself as a *certain* explanation of sensible appearances, for it cannot render accessible to the senses the reality underneath those appearances. It is then content with proving that all our perceptions are produced *as if* the reality were what it asserts; such a theory is a hypothetical explanation (ibid., p. 8).[2]

Next Duhem suggests that if we study the history of physics we soon will discover that explanatory theories consist of two distinct parts: a representative part which proposes to classify laws and an explanatory part which proposes to take hold of the reality underlying the phenomena. According to Duhem the former is the valuable part whereas the latter is useless because the underlying realities will always escape us. He continues:

> Now, it is very far from being true that the explanatory part is the reason for the existence of the representative part, the seed from which it grew or the root which nourishes its development; actually the link between the two parts is nearly always most frail and most artificial. The descriptive part has developed on its own by the proper and autonomous methods of theoretical physics; the explanatory part has come to this fully formed organism and attached itself to it like a parasite (ibid., p. 32).

So Duhemian philosophy advocates that scientific theories should primarily be mathematical contrivances representing experimental laws and should aim neither to describe an underlying reality nor to explain sensible appearances in terms of such a reality. This kind of instrumentalism we will call *philosophical* instrumentalism.

Many physical theories of the 19th century were both mechanical and explanatory. Duhem tried to discount both these aspects in remarks such as the following:

> . . . neither in Lord Kelvin's nor in Maxwell's work has the use of mechanical models shown that fruitfulness nowadays attributed so readily to it (ibid., p. 98).

And more generally:

> Towards the end of the 19th century hypothetical theories which were offered as more or less probable explanations of phenomena were extraordinarily multiplied. The noise of their battles and the fracas of their collapse have wearied physicists and led them gradually back to the sound doctrines Newton had expressed so forcefully. Renewing the interrupted tradition, Ernst Mach has defined theoretical physics as an abstract and condensed representation of natural phenomena (ibid., p. 53).

We will ignore Duhem's extraordinary claim that Newton was a positivist. (Duhem's writings contain several incorrect statements about the philosophical views of other scientists. Another example is Rankine, who Duhem mistakenly portrays as becoming a philosophical instrumentalist.)

---

[2] The emphases are Duhem's.

A philosophical instrumentalist obviously does not need to find serious scientific difficulties in an explanatory mechanical theory before he wishes to see it abandoned. Throughout his discussion of physical theories Duhem is more concerned with philosophical arguments than with scientific difficulties.

Keeping the above in mind, it is not hard to predict that any philosophical instrumentalist will, for philosophical reasons, object to any kind of realistic kinetic theory which by means of a hidden reality, such as atoms, tries to explain thermodynamics. So it comes as no surprise to find Duhem, when discussing the relation between the kinetic theory and thermodynamics, saying that:

> The latter had reached maturity and constitutional vigour when the mechanical models and kinetic hypotheses came to bring to it the assistance it did not ask for, with which it had nothing to do, and to which it owed nothing.
>
> Hence, before attributing the discovery of a theory to the mechanical models which encumber it today, it is well to make sure that these models have really presided over or aided its birth, and that they have not come like a parasitic growth and fastened themselves on a tree already robust and full of life (ibid., p. 95).

Nowhere in this discussion does Duhem examine in detail the scientific difficulties encountered by the kinetic theory; his objections are primarily philosophical in nature.

However, it does not follow that all those who use a theory as an instrument are therefore philosophical instrumentalists. A theory which is not regarded as a true explanation but which gives adequate results for minimal effort may sometimes be preferred in applied science. A scientist who is a realist and who therefore has no philosophical objections to the idea that there can be true theories which describe and explain in terms of a hidden reality may yet make use of some theory merely as an instrument. We will call such a scientist a *methodological* instrumentalist.

In addressing the problem of determining how far certain philosophical views fostered a negative attitude to the kinetic theory it is important to realize how different are what we have called philosophical and methodological instrumentalism. Consider a claim such as the following by Gardner:

> I argue that there was a gradual transition from an instrumentalist to a realist acceptance of the kinetic theory, because of gradual increases in its predictive power, 'the testedness' of its hypotheses, 'the determinateness' of its quantities, and because of resolutions of doubts about the acceptability of its basic explanatory concepts ([1979], p. 1).

It is not denied that gradual transitions from an instrumentalist to a realistic acceptance of a theory may have sometimes taken place. Cases in point are perhaps Gardner's examples of Galileo's theory of the parabolic trajectory (initially regarded as an idealization) and Planck's quantum of energy (thought to correspond to nothing in reality). However, Galileo and Planck were realists and in both cases the transition was made within this framework. Such a transition entirely *within* the framework of realism is no more than methodological instrumentalism giving way to the recognition that a theory thought to be no more than useful, is true after all.

Transitions like this cannot take place *within* the framework of *philosophical* instrumentalism. If a scientist were a philosophical instrumentalist and he made a transition to the realistic acceptance of any scientific theory then he would have abandoned one philosophy in favour of another. Only within the framework of realism can a scientist *at one and the same time* interpret some theories realistically and others instrumentalistically. But a scientist cannot simultaneously be a realist and a philosophical instrumentalist.

The central philosophical issue, the important confrontation between realism and *philosophical* instrumentalism, is not within the compass of Gardner's argument. Emphasizing empirical considerations, he does not admit that a scientist like Duhem may have deep-seated philosophical objections to any theory whatsoever based on invisible atoms. When criticising Brush's judgement that those scientists who did suggest the kinetic theory be abandoned did so not because of scientific difficulties but because of philosophical objections, Gardner comments:

> On the contrary, it seems that Planck's, Ostwald's, and Duhem's sometime preference for a phenomenological approach was based in large part upon a perception of its rival's failure to make true predictions (ibid., p. 14).

But it is a mistake to think that Duhem would have accepted a realistic interpretion of the kinetic theory if it had made truer predictions. In spite of the later excellent predictions by the theory—such as those relating to Brownian movement—Duhem never made the transition from philosophical instrumentalism to realism and he could not have done so without changing his philosophy of science.

In short, some of Gardner's arguments may be helpful when attempting to explain why a realist now thinks that a theory is true whereas formerly he thought it to be no more than a useful instrument. But the arguments are not adequate when dealing with philosophical instrumentalism which is opposed to realism in general and so, not surprisingly, to the kinetic theory in particular.

3 POSITIVISM

Positivism has its origins in the French Enlightenment and it was the philosopher Auguste Comte (1798–1857) who extended and systematised its kind of thinking. He put forward the idea that man experiences three stages of intellectual development: a primitive theological stage, a metaphysical stage, and a final positive stage. During these three stages men will only accept respectively: theological explanations, metaphysical explanations and finally descriptions of experiences. Gradually the misguided idea percolated that history decreed that eventually science would contain no metaphysical elements, included in which were 'unobservable entities'.

One who had a more direct influence on science than Comte, was Mach (1838–1916). According to him only sense perceptions are immediate and the world is made up solely of such sensations the components of which are called elements. Science can do no more than construct relations between elements and in Mach's world there is no underlying reality which can help explain what is sensed. So it is that Mach claims that:

> The assertion, then, is correct that the world consists only of our sensations. In which case we have knowledge *only* of our sensations . . .

and

> . . . there is no rift between the psychical and physical, no inside and outside, no 'sensation' to which an external 'thing' different from sensation, corresponds ([1886], p. 12, p. 310).

Mach's *philosophy* was introduced, first and foremost, as a solution to certain philosophical problems. His monism solves at one stroke the mind body problem and the problem of appearance and reality. However, radical solutions have a habit of exacting their toll by producing problems in other areas. Realist science now becomes a problem, because it accepts the independent existence of 'things' to which concepts of science may refer. The favourable outcome of scientific hypotheses referring to, say, atoms then can be explained by realist science in terms of their truth-content. Positivism has no way in which to explain how scientific theories manage to predict correctly phenomena still to be observed, and therein lies its greatest weakness. Hence, any scientific difficulties which discredit realist theories are welcome to positivists.

Mach's first outright public rejection of mechanical theories, such as the kinetic theory, came in 1872 as he himself tells us:

> In this pamphlet, which appeared in 1872, I made the first attempt to give an adequate exposition of my epistemological standpoint . . . In it both every *metaphysical* and every one-sided *mechanical* view of physics were kept away, and an arrangement, according to the principles of economy of thought, of facts—of what is ascertained by the senses—was recommended. The investigation of the dependence of phenomena on one another was pointed out as the aim of natural science ([1872], p. 9).[3]

In the 1872 work itself Mach mentions that the views expressed therein were already held by him since 1862. It is important to realise that Mach's views opposing realism in general and explanatory mechanical theories in particular, became firmly established during a period when the kinetic theory, as we shall see, was making good progress. That is, positivism in science came *first* and then came the campaign to have the kinetic theory abandoned.

---

[3] Mach's emphases. Quoted from the preface of the 1909 edition.

Also in the 1872 work, when dealing with, among others, the views of Clausius and Thomson, Mach states that:

> . . . the view that physical phenomena can be reduced to processes of motion and equilibrium of molecules is so universally spread that, at the present time, one can only let people know one's convictions are opposed to it, with caution, guardedly, and at the risk of rousing the opinion that one is not up to date and has not grasped the trend of modern culture (ibid., p. 39).

In spite of the *scientific* success of mechanical theories, and at that time particularly the mechanical theory of heat, Mach expresses philosophical convictions which run counter to this trend. According to his monism there is no reality hidden behind appearances and so there cannot be unobservable entities the motion of which is heat. So he says: 'If, then, we are astonished at the discovery that heat is motion, we are astonished at something which has never been discovered.' What, by the way of fiction, we imagine behind the appearances matters not in the least. It is quite irrelevant for scientific purposes whether we think of heat as motion or as a substance.

> One thing we maintain, and that is, that in the investigation of nature, we have to deal only with knowledge of the connections of appearances with one another. What we represent to ourselves behind the appearances exists *only* in our understanding, and has for us only the value of a *memoria technica* or a formula, whose form, because it is arbitrary and irrelevant, varies very easily with the standpoint of our culture (ibid., p. 49).[4]

Furthermore:

> It is the result of a misconception to believe, as people do at the present time, that mechanical facts are more intelligible than other, and that they can provide the foundation for other physical facts (ibid., p. 56).

No more than Duhem's, does Mach's opposition to the kinetic theory have its origins in specific scientific difficulties. Mach's convictions are anti-metaphysical and against the idea that mechanics can help explain other physical phenomena. The atom is unobservable and therefore 'metaphysical'; also, the mechanical theory of heat explains other physical facts in terms of atoms and mechanics. So the kinetic theory violates Mach's anti-metaphysical and anti-mechanistic principles.

Mach's philosophical influence on scientists increased with the publication of his *History of Mechanics* in 1883. In it Newtonian mechanics was considered in detail as were Newton's views on absolute space, time, and motion. These views were subjected to a penetrating criticism. In as far as they are a reminder that mechanical hypotheses should not be deemed essential or ultimate, Mach's observations are valuable. However, as we saw, Mach wished to go much further, because in his philosophy there is no room for mechanical *explanations.* Many of the scientists who read his work were influenced by Mach's philosophical bias and soon those who disagreed with it, as Boltzmann did, had a battle on their hands.

---

[4] Mach's emphases.

The best known scientist who was so influenced by Mach was Einstein. Mentioning how nearly all physicists of the last century at one time saw in mechanics a firm foundation for all of physics, Einstein goes on to say:

> It was Ernst Mach who, in his *History of Mechanics* shook this dogmatic faith; this book exercised a profound influence upon me in this regard while I was a student. I see Mach's greatness in his incorruptible scepticism and independence; in my younger years, however Mach's epistemological position also influenced me very greatly, . . . a position which today appears to me to be essentially untenable (Schilpp [1949], p. 21).

Einstein goes on to mention the kinetic theory as an example of a theory to which Mach's philosophy cannot do justice.

Mach's book is not a textbook on mechanics, rather its intention is to put mechanics in its place by examining it from a philosophical point of view. In the preface to the first edition Mach says that:

> The present volume is not a treatise upon the application of the principles of mechanics. Its aim is to clear up ideas, expose the real significance of the matter, and get rid of metaphysical obscurities. ([1883], p. xiii).

One of the conclusions reached is that history teaches us that man's first explanations of the world around him were of an animistic kind; these were followed by mechanical explanations and, finally, scientists will give an account of the world in accord with Mach's positivistic philosophy. To have left animistic explanations in favour of mechanical ones was a very decided and desirable advance.

> . . . Laplace even conceived a mind competent to foretell the progress of nature for all eternity, if but the masses, their positions and initial velocities were given. In the eighteenth century this joyful overestimation of the scope of the new physico-mechanical ideas is pardonable. Indeed, it is a refreshing, noble, and elevating spectacle; and we can deeply sympathize with this expression of intellectual joy, so unique in history (ibid., p. 559).

However, in turn, we must outgrow this still metaphysical phase in which we attempt explanations, albeit in mechanical rather than animistic terms. Once we were rightly pleased with mechanical explanations.

> But now, after a century has elapsed; after our judgement has grown more sober, the world conception of the Encyclopaedists appears to us as a *mechanical mythology* in contrast to the *animistic* one of the old religions (ibid., p. 559).[5]

[5] Mach's emphases.

Mach makes relatively few comments on atomism, but he does criticise atoms because they cannot be perceived by the senses and because they are invested with properties which contradict the sensations. At best the atomic theory plays a role in physics similar to auxiliary concepts in mathematics; both are nothing but mental activities. Mach's book is not a detailed scientific analysis and criticism of theories like the kinetic theory of gases; rather, it is an attack which strikes at a deeper, philosophical level. He attempts to inculcate the idea that the search for explanations, particularly mechanical ones, in terms of 'metaphysical' entities is a phase which is passing.

Boltzmann was of the opinion that by the early 1890s too much attention was being paid to Mach's ideas:

> An almost exaggerated criticism of the methods of scientific investigation is indeed a characteristic of the present day ... a whole species of modern methodological speculations have grown up, which have been expressed by Mach in the most definite and ingenious manner. He, indeed, directly maintains that the sole object of Science is the economy of labour ([1893], p. 37).

And in 1895 Boltzmann had this to say: 'It is curious to see that in Germany, ..., at present every special theory is old fashioned, while in England interest in the Theory of Gases is still active' ([1895], p. 413). Not that England was devoid of positivists. During this period Pearson, who became well known for his work in biometrics and statistics, claimed that 'sound idealism' is replacing 'the crude materialism of the older physicists'. And that 'men of science are coming to recognise that mechanism is not at the bottom of phenomena ...'. And also that 'all science is description and not explanation' ([1892], p. xiv).

At the same time, that is during the last decade of the 19th century, Ostwald and Helm were advocating their new phenomenological theory of energetics. Ostwald in particular was very explicit in expressing his appreciation of Mach's philosophy, as well as vociferous in his condemnation of atomism. However, Ostwald had a tendency to identify the postulating of mechanical hypotheses with realism and the denial of atomism with positivism. Boltzmann pointed out that Ostwald's energetics resulted from a misunderstanding of Mach's ideas. According to Mach only sensations exist and physical entities such as atoms, molecules and similar particles, do not. Boltzmann goes on:

> Ostwald understood only one half of his proposition, namely that atoms did not exist; at once he asked: what then does exist? To this his answer was that energy existed. In my view this answer is quite opposed to Mach's outlook, for which energy as much as matter must be regarded as a symbolic expression of certain relations between perceptions and of certain equations amongst the given phenomena (McGuinness [1974], p. 176).

Evidently Ostwald was philosophically confused. He was neither a positivist nor an instrumentalist, but one who attempted to construct a rival realist theory. Mechanical explanations making use of matter in the form of 'metaphysical' atoms, were to be supplanted by explanations in terms of energy.

Planck was another who had come under the spell of Mach. Yet he too was trying to find an explanation for the second law of thermodynamics. Although he accepted the second law as an empirically established fact he hoped to explain it as a consequence of radiation damping in which irreversibility was to be explained by the conversion of incident plane electromagnetic waves to outgoing spherical waves (Clark [1976], p. 90). Although influenced by the anti-mechanistic arguments of the positivists, Planck never became a thoroughgoing philosophical positivist.

It is well known how eventually Planck turned against Mach and accepted the atomic view of nature, but at this stage he thought mechanical and atomic theories old fashioned. Although in 1896 Planck severely criticised energetics and acknowledged that thanks to the mechanical view science had produced great results in the past he, significantly, adds that it is not his intention to defend the mechanical view of nature (Planck [1896], p. 73).

So it is clear that positivism and similar philosophies did not have their *origin* in specific scientific difficulties associated with the kinetic theory of gases, or with any other scientific theory. These philosophies were anti-metaphysical and anti-mechanical in character and so clashed with the principles on which the kinetic theory based its explanations. Furthermore, philosophical positivism and instrumentalism looked on the problems raised by the kinetic theory as pseudo-problems and not as genuine problems in need of solution. It was the realists who continued to grapple with these problems and, eventually, solved them.

4 THE KINETIC THEORY AND THERMODYNAMICS

Thermodynamics, although accepted by all 19th century scientists as a valuable theory, was not seen in the same light by all of them. The realists believed that the kinetic theory was a promising attempt at explaining thermodynamics. On the other hand those who did not believe in explanations for philosophical reasons considered the kinetic theory to be superfluous or even injurious and they wished to retain only phenomenological thermodynamics.

If thermodynamics is taken to be a theory of macroscopic phenomena then the kinetic theory is a theory which at deeper levels purports to be a true description of an underlying reality and also an explanation of thermodynamic phenomena and laws. Once the hypotheses of the kinetic theory have been laid down, thermodynamics will follow as a deduction from the kinetic theory. Whereas the thermodynamics which follows from the first and second laws of thermodynamics is called classical or phenomenological thermodynamics, the thermodynamics which is deduced from the kinetic theory is known as statistical thermodynamics.

If the kinetic theory makes predictions which are contradicted by true experimental data then the kinetic theory stands falsified and will be in need of revision. But will the kinetic theory also be falsified if statistical thermodynamics, which follows from it, is not in agreement with classical thermodynamics? The answer is: only if classical thermodynamics is true.

There is always the possibility that classical thermodynamics is not entirely true and that the kinetic theory as well as explaining it, corrects it. This is exactly what did happen. Statistical and classical thermodynamics differed in certain respects and it eventually turned out that the phenomenological second law, and so classical thermodynamics was not entirely true. This second law is violated by statistical fluctuations and it was the kinetic theory which explained, and statistical thermodynamics which corrected, classical thermodynamics; although this was not generally accepted until after the work of Einstein and Perrin on Brownian movement.

By the late 1860's Maxwell was convinced that from the point of view of the kinetic theory, the real character of the second law is statistical. Somewhat later Boltzmann agreed that the statistical interpretation of the second law contradicts the classical second law, because the latter brooks no spontaneous entropy fluctuations. Before the kinetic theory could be seen to correct the classical second law the existence of fluctuations would have to be demonstrated. However, both Maxwell and Boltzmann believed it infinitely improbable that *observable* fluctuations do occur.

Many were critical of a theory which produced a downright contradiction of the unconditional irreversibility which follows from the axiomatic second law. Planck, as late as 1897, comments:

> Boltzmann's theorems are an unsuccessful attempt to show how a system governed by conservative interactions can proceed irreversibly to a final state of thermodynamic equilibrium. (Clark [1976], p. 81).

Boltzmann was well aware that he could point to nothing which demonstrated the need for a statistical explanation of the second law as provided for by the kinetic theory. This put him on the defensive and he could do little more than say that the dogma of a self-sufficient phenomenology is to be avoided and that: 'No one who has fallen in love with the molecular theory will approve of its being given up completely' ([1895] & [1898], p. 445).

So, not only does the kinetic theory attempt to explain phenomenological thermodynamics in terms of unobservable entities, it also claims to correct the second law by postulating the existence of statistical fluctuations which, alas, are thought to be as unobservable as the entities themselves. Such a theory may be acceptable within a circle of devotees who have fallen in love with it, but it was obvious to the phenomenologists that such love was blind to all but metaphysical visions. To them the molecular theory seemed mere metaphysical folly.

The irony of it is that one of the kinetic theory's finest acheivements—the explanation of the second law—could with impunity be considered a failure by the opponents of this theory. However this alleged failure was turned into a resounding success when eventually the realisation came that Brownian movement bore visible witness to precisely the kind of statistical fluctuations predicted to exist by the kinetic theory.

From the above it follows that we disagree with Clark who sees the kinetic theory and thermodynamics as two quite distinct and competing research programmes. We view the kinetic theory and thermodynamics as explanans and explanandum respectively and therefore maintain they cannot be scientific rivals. The scientific problem facing the realists was to decide whether statistical thermodynamics corrected phenomenological thermodynamics or whether the latter falsified the former. Genuine scientific rivalry could only occur if there was more than one proposed explanans of thermodynamics. Alternatives to the kinetic theory as an explanans were proposed: Plancks radiation damping and energetics. However they never became serious scientific rivals to the kinetic theory because they never became scientifically successful theories.

On the other hand, the instrumentalists and positivists who for philosophical reasons did not believe in explanations thought the whole statistical approach of the kinetic theory misguided and believed phenomenological thermodynamics to be self-sufficient; a dogma to which Boltzmann, as we saw, objected. This gave rise to *philosophical* rivalry between atomism and phenomenological thermodynamics and this, contrary to what Clark holds, exerted an important influence on the criteria of choice of theories. To complicate matters further, some were only pseudo-positivists who, like Ostwald, confused anti-atomism with positivism. So whereas there was little serious scientific competition for the role of explanans of thermodynamics, there was a great deal of philosophical rivalry between atomism and phenomenological thermodynamics, as well as a certain amount of philosophical confusion. Philosophical considerations played a larger part in the opposition to the kinetic theory than did scientific difficulties.

5 THE SPECIFIC HEAT ANOMALY

Some of the difficulties encountered by the kinetic theory were the 'reversibility paradox', the 'recurrence paradox' and problems relating to the solution of the transport equations. However, the single most protracted and serious scientific difficulty was the failure to explain certain ratios of specific heats. So, if *scientific* difficulties gave rise to positivistic philosophies of science then one would have expected this failure to have played a central part in turning adherents of realism into positivists. To see what actually happened we will take a detailed look at what has become known as the anomaly of the ratio of specific heats.

Right from the start of the kinetic theory's history there was difficulty in reconciling the results obtained by the theory with the fundamental theorem of equipartition of energy. The classical theorem of equipartition states that the total energy is distributed equally between all the n mechanical degrees of freedom. For this to hold true the ratio of the specific heats will be given by

$$\gamma = C_p/C_v = (2+n)/n$$

where C_p and C_v are the specific heat of a gas at constant pressure and volume respectively.

Initially the gas particles were assumed to be perfectly smooth and rigid spheres having only the three degrees of translational freedom, that is neither rotation nor vibration were taken into consideration. For such particles, which would behave as material points, $\gamma = 5/3 = 1{\cdot}67$.

As early as 1853 Rankine, referring to the work of Herapath and Waterston, mentions that the kinetic hypothesis gives $\gamma = 1{\cdot}67$, whereas experiments show that for 'a mass of sensibly perfect gas' it is only about 1·4 ([1853], p. 321). And in 1857 Clausius deals with the discrepancy simply by saying that '. . . at every impact of two bodies, unless the same happens to be central and rectilinear, a rotary as well as a translatory motion ensues' ([1857], p. 109). However, Clausius does not show why, when introducing rotation, γ should equal 1·4. Rather an additional three degrees of freedom would have made $n=6$ and so $\gamma = 1{\cdot}33$.

Maxwell published his first paper on the kinetic theory, in 1860, because:

> So many of the properties of matter, especially in the gaseous form, can be deduced from the hypothesis that their minute parts are in rapid motion, the velocity increasing with the temperature, that the precise nature of this motion becomes a subject of rational curiosity ([1860], p. 19).

Maxwell assumes that for real particles '. . . the *vis viva* of rotation will be equal to that of translation'. Hence, theoretically $\gamma = 1{\cdot}33$, while the experimental result quoted by Maxwell is 1·408. Furthermore, by now spectroscopic evidence has shown that particles appear to have internal structure, thus adding vibrational degrees of freedom. Maxwell concludes in the 1860 British Association report (p. 16) that:

> This result of the dynamical theory, being at variance with the experiment, overturns the whole hypothesis, however satisfactory the other results may be.

So, in spite of other good results, Maxwell explicitly accepts that the hypothesis, according to which atoms behave as hard elastic spheres, has been falsified. However as Brush points out:

> This conclusion is often regarded as a 'failure' or 'paradox' of the kinetic theory although strictly speaking it is only a refutation of this particular *hypothesis* about molecules; even in the middle of the 19th century scientists did not believe that

> molecules are accurately represented by such a hypothesis. It is not the *failure* of this naive hypothesis that is surprising, but rather the *success* of the same hypothesis when used to compute many other properties of gases, there lies the paradox ([1976], p. 73).[6]

During the next decade Maxwell turns to a new theory introducing an inverse fifth power law in which molecules repel each other at a distance, rather than act by impact. At the same time Thomson, who never liked the idea of perfectly rigid particles, put forward his theory of vortex atoms. Nevertheless, none of the scientists involved doubted the existence of molecules they only wondered about their real nature. Furthermore, in 1870 Thomson published a paper describing four independent methods which lead to .similar estimates of the diameters of molecules; so something even was known about their magnitude (Brush [1976], p. 76).

In 1873 van der Waals completed a dissertation on the continuity of gaseous and liquid states. He was quite convinced of the real existence of molecules, never regarding them as a figment of the imagination, nor even as mere centres of force (Clark [1976], p. 58). His molecular theory contributed greatly to what was a very fruitful period in Dutch science. Brush describes it as follows:

> In addition to his own work, van der Waals was a stimulating influence on the development of physical sciences in the Netherlands in the last part of the 19th century. A significant part of the progress in fields such as molecular theory, physics of fluids, and low temperature physics was accomplished by Dutch scientists during the period 1870–1915 . . . ([1976], p. 254).

In his 1875 lecture to the Chemical Society Maxwell states that the statistical approach used in the kinetic theory although

> . . . somewhat difficult, does not appear faulty. On the physical side, however, it leads to consequences, some of which, being manifestly true, seem to indicate that the hypotheses are well chosen, while others seem to be irreconcilable with known experimental results, so that we are compelled to admit that something essential to the complete statement of the physical theory of molecular encounters must have hitherto escaped us ([1875], p. 499).

Whatever was missing, its absence caused the problem of the ratio of specific heats to remain unsolved. It had become progressively clearer to Maxwell just how deep-seated this problem was, at least to those who accepted that atoms truely exist. So it was that Maxwell says: 'And here we are brought face to face with the greatest difficulty which the molecular theory has yet encountered . . .'. He continues:

> If we suppose that molecules are atoms—mere material points, incapable of rotary energy or internal motion then n is 3 . . . and the ratio of the specific heats is 1·66,

---

[6] Brush's emphases.

> which is too great for any real gas. But we learn from the spectroscope that a molecule can execute vibrations of constant period. It cannot therefore be a mere point, but a system capable of changing form. Such a system cannot have less than 6 variables. This would make the greatest value of the ratio of the specific heats 1·33, which is too small for hydrogen, oxygen, nitrogen, carbonic oxide, nitrous oxide, and hydrochloric acid (ibid., p. 504).

For a clear understanding of what is to follow it is important to realize that the specific heats ratio anomaly had two parts to it. The initial problem, already raised by Rankine in 1853, was to explain the experimental value $\gamma = 1{\cdot}4$, because a theory assuming only translation gives 1·67 and one assuming both translation and rotation gives 1·33. Then there is the *additional* problem of allowing for further, possibly infinitely many, vibrational degrees of freedom each presumably demanding its equal share of the total energy and thus, in theory, causing γ to tend to 1.

However, before concluding his lecture on the kinetic theory Maxwell also says '... I must not forget to remind you of the numerous facts which it satisfactorily explains'. According to him the explanation of the various laws which express the relations between volume, pressure, and temperature, may be regarded as real and complete. 'But the most important result of these inquiries is a more distinct conception of thermal phenomena.' By this Maxwell means that a great deal has been learned about what really underlies the transport phenomena.

None of the gases mentioned by Maxwell is monatomic. At this stage γ was not known experimentally for any monatomic gases and so could not be compared with the theoretical value of 1·67 for an atom behaving like a material point. Kundt and Warburg realised the importance of performing experiments on a monatomic gas and Bayer suggested they try mercury gas. They reported their findings in 1876 ([1876], p. 353).

In their experiments they compared the velocity of sound in mercury gas with that in air and concluded that the specific heats ratio of mercury gas is 1·186 times that of air, which has the known ratio 1·405. Hence the ratio for mercury gas is $1{\cdot}186 \times 1{\cdot}405 = 1{\cdot}666$. This is exactly the figure predicted by the kinetic theory!

Kundt and Warburg were aware that they had done nothing to solve the problem of why diatomic gases have a ratio of about 1·41 instead of 1·33 or less. However they considered that they had found grounds for accepting that there was nothing wrong with the principles of the kinetic theory, but only with the manner in which it had been applied. Whatever the problem with diatomic molecules, they had shown '... that with respect to its mechanical and thermal properties the mercury gas molecule behaves remarkably like a material point.' (ibid. p. 356)[7].

---

[7] '... dass das Molekül des Quecksilbergases in Bezug auf seine mechanischen und thermischen Eigenschaften sich merklich wie ein materieller Punkt verhält.

As a consequence of the work of Kundt and Warburg an interesting suggestion was made by Bosanquet at Oxford in 1877. He put forward a conjecture showing how to explain the anomalous ratio of 1·4 ([1877], p. 271). True, a diatomic molecule will generally have six degrees of freedom and so a ratio of 1·33; but imagine a molecule of such a shape that only with respect to one of its axes it is a figure of revolution, for example a cylinder or a dumb-bell. If an addition the molecule is perfectly smooth and does not vibrate then it will have only *five* degrees of freedom. In this case $\gamma = 7/5 = 1{\cdot}4$. As Bosanquet points out, if the result of Kundt and Warburg for monatomic gases is worth anything then it can simply enough be extended to diatomic molecules. In fact:

> The relation is so obvious it is impossible to suppose that it has not occurred before to the eminent men who have dealt with the subject. (ibid. p. 276).

Unknown to Bosanquet a little earlier Boltzmann already had come to the same conclusion about diatomic molecules (Boltzmann [1877], p. 175).

So, further progress was being made in solving the first part of the anomaly. At the same time Boltzmann and Bosanquet explicitly accepted that the spectrum analysis shows that atoms and molecules cannot be absolutely rigid bodies. They further realised that unless the vibrational energy involved is very small there can be no agreement with experimental results because an appreciable amount of energy in the vibrational mode would cause the ratio of the specific heats to fall noticeably below 1·4.

Maxwell grants that some progress has been made. The values of the ratios can be explained by assuming that molecules of some gases are like particles with three, five, and six degrees of freedom '. . . becasue these values are in striking agreement with the phenomena of the three groups of gases.' ([1877], p. 245). Nevertheless he keeps on insisting that real particles cannot be perfectly rigid and so they will have an infinite number of vibrational degrees of freedom. By now Maxwell is also convinced that the same objection also applies to the vortex atoms of Thomson.

Maxwell thinks that, perhaps, the spectroscopic results might be reconciled with what we know about specific heats if we were to suppose that at ordinary temperatures the collisions between molecules are not severe enough to produce any internal vibrations so that these will only occur at very high temperatures. At once Maxwell rejects this, his near-prophetic glimpse, because: '. . . the fixed position of the bright lines in a gas shows the vibrations are isochronous . . .' (ibid., p. 246). While far from rejecting the kinetic theory Maxwell has this to say about its results:

> Some of these, no doubt, are very satisfactory to us in our present state of opinion about the constitution of bodies, but there are others which are likely to startle us out of our complacency, and perhaps ultimately drive us out of all the hypotheses in which we have hitherto found refuge into that state of thoroughly

conscious ignorance which is the prelude to every real advance in knowledge (ibid., p. 245).

Unfortunately Maxwell died in 1879. However, his idea that a variation of specific heat with temperature might have implications for the partition of energy was an interesting one. Boltzmann, and others, had assumed that during conduction of heat in a gas of polyatomic molecules the energy of translation, as compared with other types of energy, is predominately involved; this was taken to be so in order to avoid disagreement between certain theoretical and experimental results. Wüllner thought otherwise and was of the opinion that the discrepancy was only apparent, being due to the fact that the viscosity and the specific heat values, both of which entered into the equation of heat conduction, had not been determined at the same temperature.

Wüllner was led to perform experiments in which the specific heats ratio of a polyatomic gas was determined at different temperatures. He was able to show that in the cases where the specific heat increases with the temperature, all the increase is due to work done in the molecule ([1878], p. 339). Furthermore he reconciled the theoretical and empirical results relating to heat conduction and in doing so showed that Boltzmann's assumption was unnecessary.

Further researches designed to obtain additional information on ratios of specific heats were undertaken by Strecker in 1881 ([1881], p. 20). His purpose was to examine other diatomic gases to see if for these too the ratio is 1·4 as it is for O_2, H_2, N_2, CO, NO, and HCl. The gases examined by Strecker were those of chlorine, bromine and iodine. The ratio for each of these turned out to be 1·33, and so here for the first time is experimental evidence for a molecule with six degrees of freedom, and Boltzmann interprets it as such ([1881], p. 544).

A year later Strecker published another paper on the specific heat ratio of diatomic gases. He found that for HCl, HBr, and HI, the ratios are close to 1·4, but for ICl and for IBr close to 1·33 ([1882], p. 85). Again Boltzmann comments and it is clear that by now his conviction has grown firm that there is much truth in the hypothesis that the heat ratios 1·67, 1·40, and 1·33 relate to molecules behaving as rigid smooth spheres, rigid smooth figures of revolution, and general rigid smooth figures respectively ([1883], p. 309). Why, otherwise, should the same three ratios keep turning up? Equally important 1·67 and 1·40 are the *only* known experimental values greater than 1·33.

Boltzmann now feels sure that at least for those gases for which the ratio appears to be independent of temperature, and those seem to be all gases consisting of monatomic and diatomic molecules, there can be hardly any doubt that the molecules behave as smooth rigid bodies. Boltzmann also expressed the wish that more experimental work be undertaken, particularly on the heat ratio of gases consisting of polyatomic molecules because these ratios seemed to vary considerably with temperature. So up to 1883, the year of publication of Mach's *History of Mechanics*, progress was being made because the first part of the specific heats ratio anomaly now was largely solved; certainly to the satisfaction of Boltzmann who, after Maxwell's death, had become the leading figure in matters relating to the kinetic theory.

The successful solution of the first part of the anomaly depended on the validity of the theorem of equipartition of energy. So the second part assumed greater importance, not only because it still awaited a solution, but also because now the question could be asked: Why, if the theorem is valid, does it not seem to apply equally well to vibrational degrees of freedom?

Boltzmann's attitude to the anomaly was the right one. He believed that the explanation, as far as it had gone, was approximately true. The complete explanation, when found, would not be too different. This eventually proved to be the case.

Of course, under certain circumstances molecules behave very much like perfectly rigid bodies not because they *are* such, but because of quantum effects. Nevertheless it *is* the way they behave and so true explanations can be based on this behaviour, for example, the explanation of the specific heats ratios. What in turn needs explaining is *why* they behave like this, but that is another story, that of quantum mechanics.

Although Thomson and others were right to keep on insisting that real particles cannot be perfectly rigid and therefore must have vibrational degrees of freedom, explanations making use of molecules which behave like rigid bodies should not be dismissed for the sole reason that atoms and molecules are not *ultimately* rigid bodies. The claim that an explanation is ultimate is obviously not a part of the explanation; rather it is a claim about the explanation. To claim to have provided an explanation in terms of rigid bodies is very different from the claim to have found an ultimate explanation in terms of rigid bodies. Only if we insist that the only genuine explanations are *ultimate* explanations, that is if we embrace essentialism, will we reject the kinetic theory's explanation of the ratios of specific heats for the reason that atoms and molecules are not ultimately rigid bodies.

There only seemed to be two possible ways of solving the second part of the anomaly. Either assume the equipartition theorem has unrestricted validity and show how it does apply to vibrational degrees of freedom after all, or restrict the validity of the theorem. The unrestricted version had by now become known as 'the Maxwell–Boltzmann doctrine' and as 'the Boltzmann theorem'.

In 1884 Thomson gave a presidential address to the British Association in which he made it clear that he continued to oppose Boltzmann's particular kinetic theory although he accepted some kinetic theory of gases as 'a surest article of scientific faith'. Thomson preferred his vortex atoms which he believed would only have a limited number of degrees of freedom. Furthermore he did not believe Boltzmann's theorem to be true. He accepted equipartition of energy between the three translational degrees of freedom, but thought that '. . . the average kinetic energy of rotation bears a constant ratio to the average energy of translation.' ([1900], p. 486).

During the last decade of the 19th century Boltzmann's theorem and other aspects of Maxwell's treatment of the kinetic theory were subjected to intensive theoretical examination by many of Britain's major scientists. The British Association formed committees, Thomson produced 'test cases' and Tait got into a controversy with Boltzmann about it all.

In 1891 the British Association announced: 'It has now been proved beyond doubt that the theorem is not valid . . .', only to admit in 1894, in a report presented by Bryan, that a mistake had been made and it now wished to withdraw the earlier statement. Boltzmann was present at the 1894 Association meeting and there was a 'Brilliant discussion which had to be curtailed for want of time.' It was clearly realized that if, after all, the unrestricted equipartition theorem was true, then the problem of the ratios of specific heats would again be an acute one. The eventual outcome was such that Bryan hurriedly wrote to *Nature*:

> In the discussion which followed . . . Prof. Boltzmann made some remarks which appear to have been interpreted in a different light to what he doubtless intended. . . . I feel sure that Prof. Boltzmann will be much astonished to learn that his statements are now widely circulated and quoted as being an authoritative admission that the Kinetic Theory of Gases is nothing more than a purely mathematical investigation, the results of which are not in accord with physical phenomena; in short a mere useless mathematical plaything ([1894], p. 31).

In other words Boltzmann is accused of becoming a methodological instrumentalist, or worse. Bryan also mentions, having had several conversations with Boltzmann both during and after the meeting, that Boltzmann's '. . . views are not nearly so pessimistic as the opponents of the kinetic theory would wish to maintain.'

We have now followed through the kinetic theory's most serious scientific difficulty. As is clear from Bryan's remarks, as late as 1894 neither the major British participants in the development of the theory nor Boltzmann had ceased to treat it as an attempt at a true description and explanation. Far less should they be accused of having succumbed to philosophical instrumentalism or positivism. Furthermore, there is no evidence to suggest that the British kineticists later changed their minds. It has been claimed that Boltzmann did so but we will show that it is doubtful that he ever became a philosophical positivist. We may then conclude that if scientific difficulties did not force positivism on those who were most intimately aware of these difficulties then the reason for the rise of positivism *in science* must be sought elsewhere. In earlier sections we have already drawn attention to what we believe to be the reasons for this rise, namely the prior establishment of *general* positivism.

It is to be noted that the specific heats ratio anomaly continued to receive a great deal of attention from when it was first raised by Rankine in 1853. This hardly bears out Clark's statement that until 1880: 'The anomaly of the ratio of the specific heats was listed but largely ignored . . .' ([1976] p. 62).

Furthermore, Clark claims that the kinetic research programme degenerated after 1880 (ibid., p. 44). But even after 1880 general progress continued to be made, for instance by Gibbs and by the Dutch physicists; and the work of Strecker on the specific heats ratio came after this date. However the kinetic theory had done its main work. It now had to wait for experimental support for its statistical second law, which did not come until early in the 20th century; and it had to wait for quantum mechanics to solve the specific heats ratio anomaly. This waiting period in no way detracts from its achievements and any thoughts of degeneration were but subjective judgements by those opposed to it on philosophical grounds. We wholeheartedly endorse Clark's statement that: 'The theory is now regarded (rightly) as one of the greatest achievements of 19th century physics.' (ibid., p. 42). And it is not to the credit of nineteenth century positivists that they did not realize this.

6 BOLTZMANN AND POSITIVISM

Misinterpretation of Boltzmann's position began early and has persisted. In the previous section we saw how word got around after the 1894 British Association meeting that Boltzmann had given up his realism in regard to the kinetic theory and atomism. Eighty years later the same claim is still made in a book of the Vienna Circle Collection consisting of selected writings by Boltzmann. In the foreword de Groot has this to say: 'The mechanical models which he [Boltzmann] used did not need to be taken as "reality"; he did not assert that a gas consisted "really" of atoms and that the atoms "really" interact.' (McGuinness [1974], p. xi).

My thesis will be that Boltzmann defended a philosophically sophisticated form of realism while rejecting other naively optimistic realist views. His rejection of the latter has been misinterpreted as a rejection of realism. The *moderate* realism defended by Boltzmann accepts that scientific theories *aim* to be true descriptions of reality and can be used for explanatory purposes. Dogmatic realism claims that some theories can be known to be *certain* truth. Essentialism makes the additional claim that some of these theories are ultimate truths based on the '*essential*' nature of things and so cannot be further explained (Popper [1963], p. 103 and Musgrave [1977], p. 740).

Of course, we do not mean that Boltzmann explicitly recognised the various types of realism here outlined, but rather than his philosophy of science has all the hallmarks of what we have called *moderate* realism. Furthermore we do not disagree with the claim that Boltzmann moved in a direction from a more dogmatic to a more moderate type of realism. Elkana, for instance, quotes Boltzmann as mentioning such a general trend. In 1889 Boltzmann said, referring back to the 1860s:

> Eine der wichtigsten Fragen ... war die nach der Zusammensetzung der Materie. Sie ist es wohl auch noch heute; nur das man die Fragestellung etwas anders stilisiert hat. Während man damals die letzten Elemente des Seienden der Materie selbst suchte, so fragte man heute, aus welchen einfachen Elementen man die geistigen Bilder zusammensetzen muss, um die beste Übereinstimmung mit den Erscheinungen zu erzielen (Elkana [1971], p. 258).

Essentialism certainly has been left behind, but this does not imply that Boltzmann finished up as a philosophical instrumentalist or positivist. Even Elkana, when discussing Boltzmann's final position has this to say:

> Because of this new image of science there was a widespread tendency to see in Boltzmann an early logical positivist or even an instrumentalist. The commentators who claim this do not realize the depth of difference between considering a question as not any more the most important and considering it meaningless. Furthermore Boltzmann never gave up the idea that the task of natural science is to explain and not only to predict the phenomena (ibid., p. 268, n. 52).

We will now consider some of the things Boltzmann wrote during the last decade or so of his career in order to try and show how Boltzmann's views developed.

In the previous section we also saw how Bryan, who had several conversations with Boltzmann, wrote to *Nature* complaining that several of Boltzmann's statements were being misinterpreted. Bryan's assessment of the situation was confirmed in Boltzmann's reply which opens as follows:

> I propose to answer two questions:
> (1) Is the Theory of Gases a true physical theory as valuable as any other physical theory?
> (2) What can we demand from any physical theory?
> The first question I answer in the affirmative, but the second belongs not so much to ordinary physics (let us call it ortho-physics) as to what we in Germany call metaphysics ([1895], p. 413).

Boltzmann goes on to say that for a long time the theory of Boscovitch was the ideal of physics but that it is now clear that the

> ... simple conception of Boscovitch is refuted almost in every branch of science, the Theory of Gases not excepted. The assumption that the gas-molecules are aggregates of material points, in the sense of Boscovitch, does not agree with the facts. But what are they? (ibid., p. 413).

Next Boltzmann quotes Lord Salisbury as saying 'What the atom of each element is, whether it is a movement, or a thing, or a vortex, or a point having inertia, all these questions are surrounded by profound darkness.'

Boltzmann comments:

> If this be so—and hardly any physicist will contradict this—then neither the Theory of Gases nor any other physical theory can be quite a congruent account of the facts . . . Certainly, therefore, Hertz is right when he says: 'The rigour of science requires that we distinguish well the undraped figure of nature itself from the gaycoloured vesture with which we clothe it at our pleasure.' But I think the predilection for nudity would be carried too far if we were to forgo every hypothesis. Only we must not demand too much from hypotheses (ibid., p. 413).

Yet none of this meant that Boltzmann was departing from realism in regard to the kinetic theory and atomism, because he continues:

> Every hypothesis must derive indubitable results from mechanically well-defined assumptions by mathematically correct methods. If the results agree with a large series of facts, we must be content, even if the true nature of the facts is not revealed in every respect. No one hypothesis has hitherto attained this last end, the Theory of Gases not excepted. But this theory agrees in so many respects with the facts, that we can hardly doubt that in gases certain entities, the number and size of which can be roughly determined, fly about pell-mell (ibid., pp. 413, 414).

Here Boltzmann is obviously very much a realist. But what kind of realist? Boltzmann denies both that for theories to be truly scientific their truth must be established with certainty and that they must treat of the essential nature of things. That is Boltzmann was neither a dogmatic realist nor an essentialist, but rather what we have called a *moderate* realist.

Boltzmann accepted that science is not just about observations. Scientific theories make conjectural statements about all sorts of objects, some observed immediately enough, others more indirectly, as well as about how such objects act and interact. Theories, particularly new powerful ones, may contain very bold hypotheses; so bold in fact that we would be surprised if they did *not* require some modification. Once we acknowledge that scientific hypotheses are not infallible we must reject essentialism and dogmatic realism, but in no way are we thereby forced to abandon a more moderate kind of realism.

It is a feature of Boltzmann's philosophy of science that in it he dissociates himself from every kind of dogmatism. Although he was convinced of the reality of mechanical explanations he time and again points out that he does not claim them to be essential explanations. Brush mentions how that during the middle of the second half of the 19th century views were widespread that materialistic explanations were the ultimate goal of science ([1976], p. 59). Remarks in which Boltzmann distances himself from this kind of philosophical materialism, which often was associated with the idea that the ony true and exact science is mechanics, should not be misconstrued as concessions to positivism.

Furthermore, it is true that Boltzmann was in the habit of comparing thoughts and hypotheses to pictures, but this too does not imply a departure from realism. Referring to the past, he says in 1897 that: 'Nobody surely ever doubted what Hertz emphasizes in his book, namely that our thoughts are mere pictures of objects (or better signs for them) . . .' ([McGuinness], p. 225). Always having believed this did not in the past prevent Boltzmann from being a realist with respect to atoms. Why should it do so now?

A little further on in the same article Boltzmann takes up his usual position in defence of the view which pictures the atom as a mechanical system, but he does not omit to indicate that he is not opposed on principle to a different type of picture when he adds: 'I am thus the last to deny the possibility of obtaining a better picture of nature than by atomism.' (ibid., p. 227). But granting the possibility is not the same as believing that it is about to be provided. Hence: 'So long as this has not yet been done I admit the possibility but not the certainty that another world picture will replace the mechanical one.'

More is at stake here than the simple issue of realism in regard to atoms. If physical reality is not just mechanical then obviously the whole notion of atoms as simple mechanical systems may have to be revised. This Boltzmann accepted. Already in 1895 he asked of the entities flying about pell-mell in gases:

> Can it be seriously expected that they will behave exactly as aggregates of Newtonian centres of force, or as the rigid bodies of Mechanics? And how awkward is the human mind in divining the nature of things when forsaken by the analogy of what we see and touch directly? ([1895], p. 414).

The conjectures had been very bold and Boltmann was well aware that some further theoretical development would be necessary in order to have the kinetic theory of gases better describe and explain the relevant physical reality.

In 1899 Boltzmann is still taking up the same stance as before when he says to a meeting of scientists:

> I therefore present myself to you as a reactionary, one who has stayed behind and remains enthusiastic for the old classical doctrines as against the men of today; but I do not believe that I am narrow-minded or blind to the advantages of the new doctrines . . . (McGuinness [1974], p. 82).

In the same address Boltzmann mentions that the rapidly growing importance of the principle of energy has led to an attempted revolution involving the whole of physics and how this new doctrine considers the concept of energy as the only valid point of departure for inquiry into nature, while regarding every mechanical model as superfluous or even injurious. In it Boltzmann also grants that:

> . . . classical theoretical physics has led to occasional excrescences against which a reaction was necessary. Every Tom, Dick and Harry felt the call to excogitate some

> structure, some vortices, and concatenations of atoms, believing thereby to have spotted the Creator's plan once and for all (ibid., p. 93).

Farther on in the same address Boltzmann says that energeticists and phenomenologists had inferred from the small current yield of molecular theory to its decline. This theory, always thought harmful by some of them, now all of them considered superfluous. Hence: 'They all swore to annihilate it.' This worried Boltzmann. So, in 1899 Boltzmann was still concerned for the fate of the molecular theory. He admits that for the moment atomism had lost its pre-eminence, nevertheless certain valuable results have been obtained which only this theory can produce. Here Boltzmann mentions the work of van der Waals, extensions to Gibbs' theory of dissociation, and results in hydro-dynamics.

Boltzmann then goes on to claim that: 'The distinctly molecular theory of the ratio of the two specific heats of gases has likewise resumed an important role today.' He then recalls how Clausius had worked out the ratio to be 1·67 for simple spheres (which ratio did not fit any gas then known) and how Maxwell had found 1·33 for non-spherical bodies but had overlooked 1·4 for molecules symmetrical about one axis (and how 1·4 happens to be the value for the best known gases). Boltzmann continues:

> Already Kundt and Warburg had found the old Clausius value of 1·67 for mercury vapour, but because the experiment was difficult it had never been repeated and was all but forgotten. Then the same value turned up again for the ratio of specific heats in all gases newly discovered by Lord Rayleigh and Ramsay, and all other circumstances pointed to the specially simple molecular structure required by the theory, as had been the case already for mercury vapour. What influence it might have had on the history of gas theory, if Maxwell had not made this slight mistake, or if the new gases had been known as early as the time of Clausius! From the very start one would have found in the simplest gases the value that theory required for the ratio of the specific heats (ibid., p. 99).

These are hardly the utterances of a man who has had a change of heart in regard to the realistic interpretation of the atomic kinetic theory. In fact:

> All these achievements and many earlier attainments of atomic theory are absolutely unattainable by phenomenology and energetics; and I assert that a theory that achieves original insights unobtainable by other means, that is moreover supported by many facts of physics, chemistry and crystallography, such a theory should not be opposed but cultivated further (ibid., p. 99).

Nevertheless, being well aware that the present state cannot be final, Boltzmann adds:

> As regards ideas about the nature of molecules it will however be necessary to leave the widest possible room for manoeuvre. Thus one will not give up the theory of the ratio of specific heats merely because it is not yet generally applicable; for

> molecules do not behave like elastic bodies except for the simplest gases and even not for them at very high temperatures, and only as to their collisions; about their more detailed constitution which is bound to be very complicated we have as yet no indications, rather we shall have to try to find them. Parallel to atomism the further precision and discussion of the equally indispensible equations can proceed divorced from the hypothesis, without the former raising to the state of dogma its material points, or the latter its mathematical apparatus (ibid., p. 99).

Note how in spite of the specific heats ratio anomaly Boltzmann wishes to persist with the atomic theory although he, rightly, condemns dogmatic realism in regard to atoms *as material points*. He also condemns philosophical instrumentalism in regard to the mathematical apparatus. And never falling prey to narrow-minded dogmatism himself, Boltzmann in the closing remarks of his address again left open the possibility that the future may one day bring an atomic theory that is totally different from today's or even that he may be wrong and the idea of a pure continuum may afford the best picture.

Moving on to 1900 it is clear from the following that, not surprisingly, Boltzmann still rejects the claims of essentialism but, unlike the positivists, he continues to accept the notion that things exist to which theories can refer:

> Indeed, our ideas of things are never identical with the nature of things. Ideas are only pictures or rather signs that necessarily represent one-sidedly and indeed can do no more than imitate certain modes of connection of the things signified without in the least touching their essential character (ibid., p. 141).

At the same time Boltzmann is becoming more sure that mechanical explanations will require augmentation. After stating that nature would appear to us at its most comprehensible if all apparently qualitative changes were representable by motions or changes of arrangement of smallest parts, he continues:

> . . . but we cannot compel her to this, we must leave open a possibility that this will not do and that we need in addition other pictures of other changes; understandably, it is precisely the more recent developments of physics that have made it prudent to allow for this possibility (ibid., p. 143).

Recent developments in the field of electro-magnetism are proving to be of fundamental importance. In fact:

> . . . from some quarters it was attempted to replace the mechanical hegemony in the theoretical physics by an ectro-magnetic one; by trying conversely to derive the laws of mechanics from the theory of electro-magnetism (ibid., p. 143).

Now electromagnetism is making an attempt similar to that of energetics when this theory of energy became so influential '. . . as in its turn to seek hegemony in theoretical physics, attempting to make of it the science of energetics.' So there is now a three-cornered contest between explanations in terms of mechanics, electro-magnetism and energetics respectively.

In the scheme of the energeticists there was no room for atoms. However, energetics by this time was struggling and Boltzmann was convinced it was no longer a threat. On the other hand he had to agree that there were strong signs that there may be considerable truth in electro-magnetic explanations. If this turned out to be so then where would this leave realism with regard to atoms? Boltzmann was of the opinion that it might well turn out that physical reality is particulate at an even deeper level than hitherto considered and he remarks:

> . . . some very recently observed phenomena in cathode rays and electrolysis have led to the assumption that even electricity has an atomic constitution, consisting of discrete elements, namely electrons (ibid., p. 145).

In 1902 Boltzmann repeated much of the above, so we now turn to 1904. First we remind ourselves how in 1897 Boltzmann granted the possibility that a better picture might displace the mechanical one, but not the certainty that this would ever happen, and if it did, it would not be for a very long time yet. Now, in 1904, he had to admit: 'Of what I then expected after centuries or even millenia, the half has happened in seven years.' (ibid., p. 258). Not that now Boltzmann was compelled to give up his realism with regard to atoms. Rather there are discrete entities at an even lower level where purely mechanical explanations are no longer adequate.

> However, the ray of hope for a non-mechanical explanation of nature came not from energetics or phenomenology, but from an atomic theory that in its fantastic hypotheses surpasses the old atomic theory as much as their elementary structures surpass in smallness those of the old atoms. I need not mention that I mean the modern theory of electrons (ibid., p. 259).

Note that Boltzmann still dissociates himself from phenomenology. However, it now looks almost certain that the old mechanical theory will be augmented by an electro-magnetic one. Boltzmann realised this would complicate matters because he believed, as we saw, that nature would appear to us at its most comprehensible if purely mechanical.

> However, the advantage of being able to derive all mechanics from other ideas that are in any case necessary for explaining electro-magnetism would be just as great as if conversely electromagnetic phenomena could be explained mechanically. May the former succeed and my requirement of seven years ago be fulfilled! (ibid., p. 259).

Seven years ago Boltzmann had said that even though he did not despise attempts to find more general equations of which mechanical ones would only be special cases, it had to be realized that before a more comprehensive world picture could be accepted it must on no account be inferior to the mechanical one. Whereas Boltzmann always totally rejected energetics and phenomenology he thought that the new theory of electrons may lead to a genuinely more comprehensive view of physical reality and so he wished it well. This new theory is not opposed to atoms, or mechanics, or realism; in fact, Boltzmann calls it 'an atomic theory'. Furthermore, now mechanics and electromagnetism will be derivable from the same ideas.

Little further relevant material was published by Boltzmann during the remaining few years of his life, so we are now in a position to sum up Boltzmann's position during the period when he is alleged to have given up atomism and even realism.

Boltzmann's singular lack of dogmatism rightly caused him to regularly acknowledge the possibility that an atomic theory might not be the best picture to describe physical reality. Such statements are easily misinterpreted. During the period under discussion Boltzmann was genuinely uncertain as to what would constitute the most adequate description of reality at the lower levels then being reached. However, in spite of ignorance and uncertainty, so keenly felt by the non-dogmatic Boltzmann, there is little to suggest he ever became a positivist.

Boltzmann, although hardly doubting their existernce, was, like Lord Salisbury, well aware of how ignorant the then science was of the true nature of those entities which in gases 'fly about pell-mell.' His comments admitting this ignorance have often been interpreted as a rejection of realism in regard to atomism and the kinetic theory; a mistake already made at the 1894 British Association meeting.

REFERENCES

BOLTZMANN, L. [1877]: 'Ueber die Natur der Gasmoleküle' *Annalen der Physik und Chemie*, **160**, pp. 175–176.

BOLTZMANN, L. [1881]: 'Zu K. Strecker's Abhandlung: Ueber die specifische Wärme des Chlors etc.' *Annalen der Physik und Chemie*, **13**, p. 544.

BOLTZMANN, L. [1883]: 'Zu K. Strecker's Abhandlungen: Die specifische Wärme der gasförmigen zweiatomigen Verbindungen von Chlor, Brom, Jod u.s.w.' *Annalen der Physik und Chemie*, **18**, pp. 309–310.

BOLTZMANN, L. [1893]: 'On the Methods of Theoretical Physics' *Philosophical Magazine*, **36**, pp. 37–44.

BOLTZMANN, L. [1895]: 'On Certain Questions of the Theory of Gases' *Nature*, **51**, pp. 413–415.

BOLTZMANN, L. [1895] & [1898]: *Lectures on Gas Theory*. Translated by Stephen G. Brush. University of California Press, 1964.

BOSANQUET, R. H. M. [1877]: 'Notes on the Theory of Sound' *Philosophical Magazine*, **3**, pp. 271–272.

BRUSH, S. G. [1976]: *The Kind of Motion we call Heat*. North-Holland Publishing Company.

BRYAN, G. H. [1894]: 'Prof. Boltzmann and the Kinetic Theory of Gases' *Nature*, **51**, p. 31.

CLARK, P. [1976]: 'Atomism versus Thermodynamics', in *Method and Appraisal in the Physical Sciences*, editor C. Howson, pp. 41–105. Cambridge University Press.

Clausius, R. [1857]: 'On the Nature of the Motion which we call Heat' *Philosophical Magazine*, **14**, pp. 108–127

Elkana, Y. [1971]: 'Boltzmann's Scientific Research Program and its Alternatives' in *The Interaction between Science and Philosophy*, pp. 243–279. Humanities Press 1974.

Duhem, P. [1906]: *The Aim and Structure of Physical Theory*. Translated by Philip P. Wiener. Originally published by Princeton University Press. Second printing 1974; Atheneum New York.

Gardner, M. R. [1979]: 'Realism and Instrumentalism in 19th Century Atomism' *Philosophy of Science*, **46**, pp. 1–34.

Kundt, A. & Warburg, E. [1876]: 'Ueber die specifische Wärme des Quecksilbergases' *Annalen der Physik und Chemie*, **157**, pp. 353–369.

Mach, E. [1872]: *History and Root of the Principle of the Conservation of Energy*. 1911 edition. Microprint, Landmarks of Science.

Mach, E. [1883]: *The Science of Mechanics*. 1942 edition. Open Court Publishing Company.

Mach, E. [1886]: *The Analysis of Sensations*. 1914 edition. Open Court Publishing Company.

Maxwell, J. C. [1860]: 'Illustrations of the Dynamical Theory of Gases' *Philosophical Magazine*, **19**, pp. 19–32.

Maxwell, J. C. [1875]: 'On the Dynamical Evidence of the Molecular Constitution of Bodies' *Journal of the Chemical Society*, **13**, pp. 493–508.

Maxwell, J. C. [1877]: 'The Kinetic Theory of Gases' *Nature*, **16**, pp. 242–246.

Maxwell, J. C. [1878]: 'Tait's "Thermodynamics" ' *Nature*, **17**, pp. 257–259.

McGuinness, B. [1974], editor: *Ludwig Boltzmann. Theoretical Physics and Philosophical Problems*. D. Reidel Publishing Company.

Musgrave, A. E. [1977]: 'Explanation, Description, and Scientific Realism' *Scientia*, **112**, pp. 727–741.

Pearson, K. [1892]: *The Grammer of Science*. 1969 edition. Peter Smith.

Planck, M. [1896]: 'Gegen die neuere Energetik' *Annalen der Physik*, **57**, pp. 72–78.

Popper, K. R. [1956]: 'Three Views concerning Human Knowledge' in *Conjectures and Refutations*. Third edition 1969, pp. 97–119. Routledge and Kegan Paul.

Rankine, W. J. M. [1853]: 'On the Mechanical Action of Heat' in *Miscellaneous Scientific Papers*. Microprint, Landmarks of Science.

Schilpp, P. A. [1949], editor: *Albert Einstein, Philosopher-Scientist*. 1959 edition. Harper Torchbook.

Strecker, K. [1881]: 'Ueber die specifische Wärme des Chlor-, des Brom- und des Jodgases' *Annalen der Physik und Chemie*, **13**, pp. 20–42.

Strecker, K. [1882]: 'Die specifischen Wärmen der gasförmigen zweiatomigen Verbindungen von Chlor, Brom und Jod unter sich und mit Wasserstoff' *Annalen der Physik und Chemie*, **17**, pp. 85–103.

Thomson, W., Lord Kelvin [1900]: 'Nineteenth Century Clouds over the Dynamical Theory of Heat and Light' in *Baltimore Lectures*, Appendix B, pp. 486–527. C. J. Clay 1904.

Wüllner, von A. [1878] 'Ueber die Abhängigkeit der specifischen Wärme der Gase bei constantem Volumen von der Temperatur und die Wärmeleitungsfähigkeit der Gase' *Annalen der Physik und Chemie*, **4**, pp. 321–341.

Brit. J. Phil. Sci. 39 (1988), 111–116 Printed in Great Britain

DISCUSSION

On the Alleged Impossibility of Inductive Probability*

Karl Popper and David Miller [1983] have argued that there is no such thing as probabilistic inductive support, as conceived of, for example, in the Bayesian theory of evidence. A little more specifically, they argue that 'all probabilistic support is purely deductive' (p. 688), and that only probabilistic countersupport (disconfirmation) could be inductive in nature. Their interesting and striking argument, which they say is 'completely devastating to the inductive interpretation of probability' (p. 688), has received considerable attention, both in the way of criticism and in the way of defense: Richard Jeffrey [1984] and I. J. Good [1984] have criticised the argument, Isaac Levi [1984] attempts to improve it, Popper and Miller [1984] defend it from Jeffrey's and Good's criticisms, M. L. G. Redhead [1985] offers a new kind of criticism, and Donald Gillies [1986] defends the argument from Redhead's objection. I will not discuss these responses here. After a brief rehearsal of the Popper–Miller argument, I will criticize it in a way different from the criticisms just cited. I will argue that the Popper–Miller argument fails to establish its conclusion on any interesting way of understanding its conclusion. This will involve a little clarification of the structures of inductive and deductive support.

According to the Bayesian theory of probabilistic inductive support, the degree to which evidence e supports a hypothesis h is given by the measure:

$$s(h/e)=p(h/e)-p(h),$$

where p is an appropriate probability measure, $p(h)$ is the prior probability of h, and $p(h/e)$ is the posterior probability of h on the evidence e ($p(h/e)=p(h\ \&\ e)/p(e)$).[1] If $s(h/e)$ is positive, then e confirms h to the degree given by the difference; disconfirmation (countersupport) and its degree are indicated by a negative $s(h/e)$ and its magnitude; and evidential neutrality is indicated by $s(h/e)=0$.

---

* I thank Elliott Sober for helpful comments on an earlier draft, and the John Simon Guggenheim Memorial Foundation, the National Science Foundation (grant no. SES-8605440), and the University of Wisconsin-Madison Graduate School Research Committee for financial support during the time this paper was written.

[1] For some interpretations of probability, it is necessary to accommodate background knowledge, b, explicitly, yielding the measure $s(h/e,b)=p(h/b\ \&\ e)-p(h/b)$. But for a subjective interpretation of probability, assumed in Bayesian confirmation theory, background knowledge is already implicitly accommodated in the subjective probability function p: $p(b)=1$, so that $p(-/b)=p(-)$.

Popper and Miller call s *probabilistic* support; what they question is whether it really measures *inductive* support.

Here is the gist of the Popper–Miller argument.[2] For any hypothesis h and evidence e, h is logically equivalent to the conjunction $(h \vee e)$ & $(h \vee -e)$. For simplicity, assume in what follows a probability p that assigns only nonextreme (not 0 and not 1) values to h, to e, and to nontautologous truth-functionally consistent truth-functional compounds of h and e. Then it is easy to see that:

$$(*) \quad s(h/e) = s(h \vee e/e) + s(h \vee -e/e).$$

It is also easy to see that $s(h \vee e/e)$ is necessarily positive and that $s(h \vee -e/e)$ is necessarily negative. Popper and Miller point out that the disjunction $h \vee e$ deductively follows from e and that it is the strongest part ('part' means *truth-functional implication*) of h that deductively follows from e. Thus, since $h \vee -e$ is the weakest part of h that in conjunction with $h \vee e$ is equivalent to h, they call $h \vee -e$ 'all of h that goes beyond e' ([1983], p. 687).[3]

Thus, when h is 'factored' into the strongest part of it that is deductively implied by e and the part (the rest) of it that 'goes beyond e', we see that e probabilistically supports only the former, and that it countersupports the latter. As Popper and Miller put it, 'that part of a hypothesis that is not deductively entailed by the evidence is always strongly countersupported by the evidence' ([1983], p. 688). The other part of the hypothesis, the part that is deductively entailed by the evidence, is the only part that is supported. They conclude that 'all probabilistic support is purely deductive' ([1983], p. 688).[4] And they point out that their argument is completely general. Indeed, if their argument is correct, 'it holds for every hypothesis h: and it holds for every evidence e, whether it [probabilistically] supports h, is independent of h, or countersupports h' ([1983], p. 688).

I wish to contest the inference from 'e probabilistically supports only that part of h that is deductively entailed by e' to 'all probabilistic support (as opposed to countersupport) is purely deductive.' I shall maintain that on a proper understanding of inductive support, even if an item X deductively entails an item Y, some aspects of X's suport of Y, may be purely *inductive* in nature. This can perhaps most readily be seen from the fact that, given that X implies Y, $s(Y/X)$ is a function wholely of $p(Y)$ (the function is $1-p(Y)$)—*where $p(Y)$ is completely independent of deductive relations between X and Y.* But I shall first adduce some more or less independent motivation for my claim.

---

[2] As somewhat reformulated and simplified by Jeffrey [1984] and Gillies [1986].

[3] Jeffrey [1984] contests this characterization of $h \vee -e$, and Popper and Miller [1984] respond. But this issue is irrelevant to my main point here, which will focus on e's deductive and inductive significances for $h \vee e$.

[4] Alternative ways of factoring h into two components are discussed in Levi [1984], Jeffrey [1984], Good [1984], and Popper and Miller [1984]. This issue also is irrelevant to my main point here, since what I say will apply to the deductive and inductive significances of e for *any* factor of h that is a deductive implication of e.

The Popper–Miller critique of the theory of probabilistic inductive support shows that the degree of e's probabilistic support or countersupport for h ($s(h/e)$) is a function (addition) of (i) $s(h \vee e/e)$, which they seem to associate with e's purely deductive support for h, and (ii) $s(h \vee -e/e)$, the degree of e's probabilistic (and inductive) countersupport for h. To help motivate my claim of the previous paragraph, consider the following two numerical examples. In the two examples, the degree of the evidence's probabilistic inductive *countersupport* for the hypothesis (item (ii) above, that is, $s(h \vee -e/e)$) is the same (I'll use upper case 'H' and 'E' for the second example, so that $s(h \vee -e/e) = s(H \vee -E/E)$). However, in the first example, the evidence *supports* the hypothesis overall, and in the second, the evidence *countersupports* the hypothesis overall (so there is a difference in that $s(h/e) > 0 > s(H/E)$). The only difference therefore must lie in what Popper and Miller seem to identify with a piece of evidence's *deductive support* for a hypothesis (item (i) above, that is, $s(h \vee e/e)$ doesn't equal $s(H \vee E/E)$). What I dispute is that this difference is a 'purely deductive' difference between the two examples: part of the difference, I claim, is inductive in nature.

EXAMPLE 1:

$$\begin{array}{ll} p(h \,\&\, e) \quad = \cdot 3 & p(-h \,\&\, e) \quad = \cdot 2 \\ p(h \,\&\, -e) = \cdot 2 & p(-h \,\&\, -e) = \cdot 3 \end{array}$$

So:

$$\begin{array}{lll} p(h/e) = \cdot 6 & p(h) = \cdot 5; & s(h/e) = \cdot 1 \\ p(h \vee e/e) = 1 & p(h \vee e) = \cdot 7; & s(h \vee e/e) = \cdot 3 \\ p(h \vee -e/e) = \cdot 6 & p(h \vee -e) = \cdot 8; & s(h \vee -e/e) = -\cdot 2 \end{array}$$

EXAMPLE 2:

$$\begin{array}{ll} p(H \,\&\, E) \quad = \cdot 3 & p(-H \,\&\, E) \quad = \cdot 2 \\ p(H \,\&\, -E) = \cdot 4 & p(-H \,\&\, -E) = \cdot 1 \end{array}$$

So:

$$\begin{array}{lll} p(H/E) = \cdot 6 & p(H) = \cdot 7; & s(H/E) = -\cdot 1 \\ p(H \vee E/E) = 1 & p(H \vee E) = \cdot 9; & s(H \vee E/E) = \cdot 1 \\ p(H \vee -E/E) = \cdot 6 & p(H \vee -E) = \cdot 8; & s(H \vee -E/E) = -\cdot 2 \end{array}$$

In these two examples, ***the relevant deductive relations among the relevant items are exactly the same.*** In particular, the forms of the relevant deductive relations between e and $h \vee e$ on the one hand, and between E and $H \vee E$ on the other, are exactly the same. To make this point even more vivid, example 2 could have been described in terms of *the very same items* h and e in terms of which example 1 was described—but with different probability and support functions, P and S, corresponding, for example, to the subjective probabilities of a different agent. But the only items involved in the examples that are capable of standing in relevant deductive relations to each other are the pieces of evidence, the hypotheses, and their truth-functional compounds. Thus, it is clear that all the deductive relations among the items in example 1 that are capable of standing in relevant deductive relations (namely, h, e, and their truth-functional compounds) are *exactly the same* as the deductive relations among the corresponding items of example 2 that are capable of standing in relevant deductive relations (namely, H, E, and their truth-functional compounds). So it seems natural to say that e's deductive support for h (or e's deductive support for $h \vee e$) is exactly the same as E's deductive support for H (or E's deductive support for $H \vee E$). Also, as noted above, e's probabilistic inductive countersupport for h (or for $h \vee -e$) is exactly the same as E's probabilistic inductive countersupport for H (or for $H \vee -E$), for they have the same s-value. But e's overall probabilistic support for h is different from E's overall probabilistic support for H. It follows that the overall probabilistic support that evidence provides for a hypothesis cannot be any function of just the evidence's deductive support and its probabilistic inductive countersupport for the hypothesis—and there must be some difference between the two examples having to do with *nondeductive* support.

All of this is consistent with (*) and the fact that evidence only probabilistically supports the part of a hypothesis that it deductively implies and probabilistically countersupports the rest of it. I believe Popper and Miller are wrong in associating $s(h \vee e/e)$ (of (*)) with 'purely deductive support' of $h \vee e$ by e. The key to disentangling these two ideas is to be careful in distinguishing between *support that is purely deductive in nature* and *support that happens to be for a deductive implication of the evidence.* This perhaps is a subtle distinction, but, as clarified and argued below, the fact that a given piece of evidence supports an item deductively implied by the evidence doesn't mean that that support is purely deductive in nature.

Popper and Miller associate a *degree* with the component of an evidence's support of a hypothesis that they call purely deductive support: $s(h \vee e/e)$ (and $s(H \vee E/E)$). But, properly understood, it seems that support that is purely deductive in nature is an 'all or nothing' affair; either the evidence fully guarantees the truth of the hypothesis (deductively implies it) or it does not fully guarantee the truth of the hypothesis (does not deductively imply it). Purely deductive support does not come in degrees.

What, then, is the significance of $s(h \vee e/e)$ (and of $s(H \vee E/E)$)? It is the *difference* between the posterior and prior probability of $h \vee e$ on e. The fact that the posterior probability is 1 is a consequence of the fact that e deductively implies $h \vee e$, *but it is this fact alone about* $s(h \vee e/e)$*—along with consequences of this fact, such as the measure's necessarily being nonnegative—that has anything to do with e's deductive support of* $h \vee e$. The *particular magnitude* of $s(h \vee e/e)$, being a *degree* and a *function partly of* $p(h \vee e)$, clearly 'goes beyond' the deductive relations between e and $h \vee e$. This naturally suggests *inductive* support. The question therefore arises of whether this aspect of $s(h \vee e/e)$ could, for additional reasons, be correctly described as representing *inductive* support of $h \vee e$ by e.

The answer quite clearly is 'yes.' Induction has to do with the proper way of learning from experience. Stated in (Bayesian) terms of subjective probability, it is the process of properly changing one's degrees of belief when confronted with evidence. Evidence confirms (supports) or disconfirms (countersupports) a hypothesis (for the relevant person) according to whether the (subjective) probability of the hypothesis is (or would be) properly changed to a higher or lower value upon learning the evidence. And the magnitude of the change naturally measures the degree of confirmation or disconfirmation. The measure $s(h \vee e/e)$ is just that magnitude in the case in question.

Thus, equation (*) should be interpreted as: the degree of e's probabilistic (and, I say, inductive) support for h is the sum of (i) the degree of *e's probabilistic inductive* support for $h \vee e$ and (ii) the degree of e's probabilistic inductive countersupport for $h \vee -e$. Item (i) here is represented in (*) as the value $s(h \vee e/e)$. This value is a *degree*; it is a *function in part of* $p(h \vee e)$; and, as this and the two examples above show, it is *not a function just of the deductive relations between e and* $h \vee e$—hence, item (i) cannot represent purely deductive support, as Popper and Miller claim. And since it measures our change in degree of belief in $h \vee e$ upon learning e, it is natural to take it as representing the probabilistic inductive aspect of e's support of $h \vee e$.

Of course, given any hypothesis h and evidence e, deductive relations between e and h can impose restrictions on $s(h/e)$. In particular, if e deductively implies h, then $s(h/e)$ has to be nonnegative. But there is more to say about the inductive significance of e for h than simply that $s(h/e) \geqslant 0$). Conversely, attending to deductive relations between evidence and hypothesis can sometimes tell us things about their probabilistic relations to one another that attending to inductive relations alone doesn't. For example, if e deductively implies h, this tells us that $p(h/e) = 1$, whereas knowing only the value of $s(h/e)$ tells us no such thing. But again, there is more to the inductive significance of e for h than that $p(h/e) = 1$: there is also the *change* in probability. And to describe this change, it is not enough to say where the probability of the hypothesis *ends up*, which is all that can be gathered from the deductive relations between evidence and hypothesis in the cases in question where the evidence implies the hypothesis. When the evidence implies the hypothesis, it is easy to overlook the purely inductive aspects of the evidence's support of the hypothesis. For if the evidence is learned, the deductive bearing of the evidence on the hypothesis settles the question of the *truth* of the hypothesis—which is usually the important question—whereas the purely inductive bearing (i.e., the value of s) does not.

Thus, I believe that Popper and Miller have failed to establish the conclusion that all probabilistic (positive) support is purely deductive support and all probabilistic inductive support is countersupport. They *have* shown that a hypothesis can be divided into two parts such that the evidence only probabilistically supports the part of the hypothesis that the evidence deductively implies, and probabilistically countersupports the rest. But if we are careful in distinguishing between the ideas of *support that is purely deductive in nature* and *support of a deductively implied hypothesis*, it is easy to see that their argument fails to establish the conclusion that all probabilistic support is purely deductive in nature.

Also, the conclusion that *is* established by Popper and Miller's argument (noted just above) is hardly 'completely devastating to the inductive interpretation of the calculus of probability.' For, as shown above, this 'weaker' (or less interesting) conclusion is perfectly compatible with the idea that whether any evidence e confirms, disconfirms, or is neutral for any hypothesis h turns on *nondeductive*—indeed *probabilistic inductive*—features of the bearing of e on h, indeed, on *both* of the two components into which Popper and Miller divide h: $h \vee e$ (which e probabilistically inductively supports to the degree $s(h \vee e/e)$) *and* $h \vee -e$ (which e probabilistically inductively countersupports to the degree $s(h \vee -e/e)$).

ELLERY EELLS
University of Wisconsin-Madison

REFERENCES

GILLIES, D. [1986]: 'In Defense of the Popper-Miller Argument', *Philosophy of Science*, 53, pp. 110–113.

GOOD, I. J. [1984]: 'The Impossibility of Inductive Probability', *Nature*, 310, p. 434.

JEFFREY, R. C. [1984]: 'The Impossibility of Inductive Probability', *Nature*, 310, p. 433.

LEVI, I. [1984]: 'The Impossibility of Inductive Probability', *Nature*, 310, p. 433.

POPPER, K. R. and MILLER, D. [1983]: 'A Proof of the Impossibility of Inductive Probability', *Nature*, 302, pp. 687–688.

POPPER, K. R. and MILLER, D. [1984]: 'The Impossibility of Inductive Probability', *Nature*, 310, p. 434.

REDHEAD, M. L. G. [1985]: 'On the Impossibility of Inductive Probability', *British Journal for the Philosophy of Science*, 36, pp. 185–191.

Brit. J. Phil. Sci. 39 (1988), 117–129 *Printed in Great Britain*

REVIEW ARTICLE
Keynes as a Methodologist*

1 *Introduction*
2 *Keynes' theory of long-term expectation and animal spirits*
3 *Keynes' Methodology compared with Popper's*
4 *Keynes on Econometrics*

1 INTRODUCTION

In 1983, the centenary of Keynes' birth, the *Cambridge Journal of Economics* organised a conference on methodological issues in Keynes' work. The present volume contains eleven of the conference papers in an extended and developed form (Chs. 2–12), together with an introduction by the editors (Ch. 1). The theme is a particularly interesting one, since Keynes started his academic life as a philosopher of science. His first piece of research was on the philosophy of probability, and it was for this that he was awarded a fellowship at King's College Cambridge. The question naturally arises as to whether Keynes' early work on probability influenced his later economics. This problem is tackled by several of the contributors, and many interesting connections are brought to light. Indeed most of the book is concerned, in one way or another, with probability. Only the last three chapters deal with exclusively economic issues. In Ch. 10, Laurence Boland discusses endogenous and exogenous variables in neoclassical and Keynesian economics, in ch. 11. Victoria Chick makes some interesting points about time and the wage-unit, while ch. 12 by Juergen Backhaus discusses the extent to which German recovery in the 1930s was consciously inspired by Keynes' ideas. These three chapters will be of particular interest to economists but, in this review, I shall concentrate on the main theme of the collection, that is to say issues about probability, induction, and the methodology of statistics. These issues will be very familiar to most readers of the *British Journal for the Philosophy of Science.* Yet it is fascinating and refreshing to see them discussed by economists in the context of economics. New light is shed on many old questions, and, altogether, this collection is much to be recommended to the philosophy of science community.

---

* Review of Tony Lawson and Hashem Pesaran (eds.) [1985]: *Keynes' Economics. Methodological Issues* Croom Helm. 265 pp. Hardback £19.95. ISBN 0-7099-1680-9.

2 KEYNES' THEORY OF LONG-TERM EXPECTATIONS AND ANIMAL SPIRITS

Chapters 2, 3 and 4 are concerned with an important part of Keynes' mature economic system: his theory of long-term expectation and animal spirits. Dow and Dow in Ch. 3 relate this to Keynes' earlier work on probability (p. 46):

> 'The first task is to go back to the origins of the concept of animal spirits in Keynes' early work on probability theory and its development in his theory of entrepreneurial behaviour.'

I agree with this suggested link, but would like to make a couple of points designed to clarify its nature. First, however, it would be as well as to explain briefly Keynes' views on long-term expectation. These are set out in the famous Ch. 12 of Keynes' [1936] General Theory. This chapter is entitled 'The State of Long-Term Expectation'. A related account is to be found in Keynes' [1937] article: 'The General Theory of Employment', originally published in *The Quarterly Journal of Economics.*

Keynes begins his [1937] treatment of the question by giving as examples of things which are uncertain in his sense of the term, [1937] p. 113: 'the price of copper and the rate of interest twenty years hence'. He then continues, [1937] p. 114:

> 'About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know. Nevertheless, the necessity for action and for decision compels us as practical men to do our best to overlook this awkward fact and to behave exactly as we should if we had behind us a good Benthamite calculation for a series of prospective advantages and disadvantages, each multiplied by its appropriate probability, waiting to be summed.
>
> How do we manage in such circumstances to behave in a manner which saves our faces as rational economic men?'

Keynes answers this last question by saying that we resort to 'a variety of techniques' of which the most important is the following, [1937] p. 14:

> 'Knowing that our own individual judgement is worthless, we endeavour to fall back on the judgement of the rest of the world which is perhaps better informed. That is, we endeavour to conform with the behaviour of the majority or the average. The psychology of a society of individuals each of whom is endeavouring to copy the others leads to what we may strictly term a *conventional* judgement.'

Keynes' point is that because of lack of information and because of the general uncertainty of the future, entrepreneurs cannot form a scientific or rational expectation, but, as they have to act, they substitute for this a conventional expectation, which then determines their investment decisions. But, just because this expectation is largely conventional, it is subject to waves of optimism or pessimism, the general state, that is, of the famous animal spirits, Keynes [1936] pp. 161–2:

> '. . . there is the instability due to the characteristic of human nature that a large proportion of our positive activities depend on spontaneous optimism rather than on a mathematical expectation, whether moral or hedonistic or economic. Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as a result of animal spirits—of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities. . . . Thus if the animal spirits are dimmed and the spontaneous optimism falters, leaving us to depend on nothing but a mathematical expectation, enterprise will fade and die;—though fears of loss may have a basis no more reasonable than hopes of profit had before.'

But how does all this relate to Keynes' early views on probability as set out in his [1921] *A Treatise on Probability?* At first the link may seem rather tenuous, since Keynes in his [1921] *Treatise* identified probability with degree of rational belief. The degree of rational belief in *a* given *b* was something which could be calculated by purely logical means and was the same for all rational men. We are a far cry here from the influence of animal spirits.

There are nonetheless passages in the [1921] *Treatise on Probability* which look forward to the [1936] *General Theory.* One such is Ch. 3 entitled: 'The measurement of probabilities'. Here Keynes considers examples like the underwriters at Lloyd's (p. 23), and concludes, [1921] Ch. 3.6 p. 29: '. . . no exercise of the practical judgement is possible, by which a numerical value can actually be given to the probability of every argument.' If, however, degrees of belief cannot be rationally calculated in every case, this leaves room for their magnitude to be determined by conventional judgements and animal spirits. It is interesting to note that Keynes' fellow Cambridge don: Sir Harold Jeffreys took a different view. Jeffreys, like Keynes, attended the lectures of the Cambridge logician W. E. Johnson, and, again like Keynes, Jeffreys developed the view of probability as degree of rational belief. Unlike Keynes, however, Jeffreys thought that for any two propositions *a* and *b*, the probability of *a* given *b* had a well-defined value (cf. Jeffreys [1939] Preface and Chapter I).

There is at least one passage in his treatment of long-term expectation in the [1936] *General Theory*, where Keynes seems to refer back to his early work on probability. Keynes writes, [1936] p. 152:

> 'Nor can we rationalise our behaviour by arguing that to a man in a state of ignorance errors in either direction are equally probable, so that there remains a mean actuarial expectation based on equi-probabilities. For it can easily be shown that the assumption of arithmetically equal probabilities based on a state of ignorance leads to absurdities.'

In fact Keynes had concluded in the *Treatise*, [1921] ch. 4.1 p. 44 that 'In order that numerical measurement may be possible, we must be given a number of *equally* probable alternatives'. In order to obtain these equally probable alternatives, Keynes appealed to the *Principle of Indifference* which states roughly that, if there is no reason for preferring any one of a number of alternatives to any other, we should assign them equal probabilities. In Ch. 4 of the [1921] *Treatise*, Keynes showed that the Principle of Indifference leads to paradoxes in a number of cases, but he still held it to be valid in others. By [1936], however, his attitude to the Principle of Indifference seems to have hardened, and he appears to reject it in all cases as leading to absurdities. This is an indication that Keynes' views on probability had actually altered between 1921 and 1936. I believe that this was indeed the case, and that the cause of the alteration is not hard to discover. It was in fact the influence of Ramsey, whom Keynes held in the very highest esteem.

Ramsey in his [1926] paper Truth and Probability subjected Keynes' views on probability to an extensive criticism. Ramsey died in 1930 at the age of only 26, and Keynes paid a tribute in Ch. 29 of his [1933] Essays in Biography to this remarkable Cambridge mathematician, economist, and philosopher. This is what Keynes says about Ramsey's treatment of probability, [1933] pp. 338–9:

> 'Ramsey argues, as against the view which I had put forward, that probability is concerned not with objective relations between propositions but (in some sense) with degrees of belief, and he succeeds in showing that the calculus of probabilities simply amounts to a set of rules for ensuring that the system of degrees of belief which we hold shall be a *consistent* system. Thus the calculus of probabilities belongs to formal logic. But the basis of our degrees of belief—or the *a priori* probabilities, as they used to be called—is part of our human outfit, perhaps given us merely by natural selection, analogous to our perceptions and our memories rather than formal logic. So far I yield to Ramsey—I think he is right. But in attempting to distinguish 'rational' degrees of belief from belief in general, he was not yet, I think, quite successful.'

We see that Keynes was prepared to yield to Ramsey on a number of points. Now it is one feature of Ramsey's approach that it can dispense with the Principle of Indifference. Ramsey remarked on this point, [1926] p. 189:

> 'To be able to turn the Principle of Indifference out of formal logic is a great advantage; for it is fairly clearly impossible to lay down purely logical conditions for its validity, as is attempted by Mr. Keynes.'

This comment may lie behind Keynes' apparently total rejection of the Principle of Indifference in 1936 which we noted above.[1]

Keynes does not however totally capitulate to Ramsey, but has doubts as to whether Ramsey has given a satisfactory account of the difference between degree of belief, and degree of rational belief. Let us examine this point in relation to Keynes theory of long-term expectation.

---

[1] For further details of Keynes on the Principle of Indifference, and of Ramsey's criticisms of Keynes see my [1988].

Keynes' original probabilities or degrees of rational belief were the same for all rational human beings. Ramsey's probabilities or degrees of subjective belief were related to a particular individual, and so could vary from individual to individual. The expectations of entrepreneurs in a given economy are, according to Keynes' later theory, somewhat intermediate between these two extremes.

Keynes does not postulate, as a strict follower of Ramsey might have done, that each entrepreneur forms his own individual expectation which differs from that of every other entrepreneur. On the contrary the entrepreneurs imitate each other so that the group comes to have more or less the same expectation. However this objective (or at least inter-subjective) expectation is not based on a rational assessment but depends on factors like the state of the animal spirits. The approach here is an example of Keynes' preference for group or aggregate concepts in his analyses of social life.

Keynes' account of expectation could be developed into a full-blown philosophy of probability in which probability would be seen not as the degree of subjective belief of an individual (as in Ramsey and de Finetti), but rather as the degree of inter-subjective belief of a group, which, through a process of social interaction, reaches a consensus. Keynes' analysis of entrepreneurs could thus be extended to groups of scientists or technologists, or to social classes within society. There are here some interesting possibilities for theoretical development.

3 KEYNES' METHODOLOGY COMPARED WITH POPPER'S

Ch. 6 by Pheby is devoted to a detailed comparison between Keynes' methodology and Popper's. This theme is also mentioned by other contributors. For example, Lawson in Ch. 7 compares Keynes' attitude to predictions to that of the Popperians, a question which he has also treated in his [1985] article in the present journal.

Keynes wrote in a famous passage, [1921] Ch. 25.3 p. 337:

> 'The peculiar virtue of prediction . . . is altogether imaginary. The number of instances examined and the analogy between them are the essential points, and the question as to whether a particular hypothesis happened to be propounded before or after their examination is quite irrelevant.'

The opposite view is that theories are better confirmed by successful predictions, than by the explanation of facts already known when the theory was invented. Confirmation theories which accept Keynes' point of view on this question are known as *logical* theories of confirmation, while confirmation theories which accord special significance to predictions or 'novel' facts are known as *historical* theories of confirmation.[2]

Lawson attributes an historical approach to confirmation to the Popperian school, and himself defends the logical or Keynesian approach. As he says (p. 123):

> 'The importance of the successful prediction of novel "facts" makes the historical context of prediction important in the falsificationist account.'

Now Lawson is undoubtedly right as regards the majority of the Popperian school, and, in particular, as regards the followers of Lakatos. It is, nevertheless, possible to develop Popper's views on corroboration in a way which is quite consistent with Keynes' *dictum* regarding predictions.

To see this let us consider the following key passage from Popper, [1959]. Appendix *ix p. 418:

> '. . . I must insist that $C(h,e)$ can be interpreted as degree of corroboration only if e is *a report on the severest tests we have been able to design.*'

If we interpret a severe test of h as an experiment or observation carried out after h has been formulated, then Popper's theory of corroboration is strongly historical, and quite the opposite of the view of Keynes quoted above. There is, however, another sense of test, in which an hypothesis h can be tested against a body of data already collected. In this sense of test, Kepler may be said to have tested out his laws against the data which had been obtained by Tycho Brake several years before. In my [1973] I allowed such a manner of speaking by defining test in the following broad fashion (p. 67).

> 'We deduce from the theory (plus, of course, certain initial conditions and perhaps other assumptions) that a certain proposition should hold exactly or at least approximately. The proposition must be such that its truth or falsity can be or *has been* established experimentally. The theory passes the test if it gives the proposition the same truth—value as that assigned to it by the experiment, and fails the test otherwise.'

From this definition, it was a straightforward matter to conclude (p. 67) that

> '. . . a theory can be corroborated just as well by experimental results established before its invention as by new results established later and suggested by the theory.'

Thus Popper's theory of corroboration can be developed in a manner fully consistent with Keynes' *dictum* on predictions.

Historical theories of confirmation have recently been criticised in an incisive fashion by Lawson himself in his [1985] and by Howson in his [1984]. These criticisms seem to me correct, but this does not lead me to accept the

---

[2] For a fuller explanation of this distinction, and an excellent critical discussion of the Popperian and Lakatosian literature on the subject, see Alan Musgrave [1974].

Bayesianism adopted by Howson, or by Lawson (following Keynes). My own preferred option is the neo-Popperian approach to confirmation sketched above.

There need not then be such a wide gap between Keynes and Popper on prediction as Lawson suggests. This observation leads us on naturally to Pheby's contribution, for Pheby's main claim is that the methodologies of Keynes and Popper are much more similar than might at first be thought. This is how Pheby summarises his view (p. 110):

> '... our considerations so far point to some substantial similarities in the methodologies adopted by Keynes and Popper. They both, in effect, combine induction and deduction, employ a predominantly problem-dependent approach, adhere to critical rationalism and an associated dislike of dogmatism, and are highly suspicious of attempts to "formalise" the social sciences.'

Pheby argues convincingly for most of the points he makes here but a doubt remains concerning the alleged similarity of Keynes and Popper on induction. Can the Bayesian inductionist Keynes really be similar on this point to the great critic of induction? My answer is 'no'. It seems to me that Keynes and Popper do differ as regards induction, and that the difference is a significant one, since, as I shall argue in the next section, it is crucial for the assessment of Keynes' critique of econometrics. My own views on this question are best approached by a consideration of Pheby's.

Pheby writes (p. 107): '... Popper's more recent developments of degrees of corroboration ... are not without their inductive features'. This might seem correct for Popper himself says, [1959] Appendix *ix p. 418:

> 'It might well be asked at the end of all this whether I have not, inadvertently, changed my creed. For it may seem that there is nothing to prevent us from calling $C(h,e)$ 'the inductive probability of h, given e'.... A benevolent inductionist critic might even congratulate me on having solved, with my C function, the age-old problem of induction *in a positive sense*—on having finally established, with my C function, the validity of inductive reasoning.
>
> My reply would be as follows. I do not object to calling $C(h,e)$ by any name whatsoever, suitable or unsuitable: ... Nor do I object—so long as it does not mislead us—to an extension (inadvertent or otherwise) of the meaning of "induction".'

Thus Popper appears to suggest that we might interpret his measure of corroboration inductively. However Popper does add immediately [1959], Appendix *ix p. 418:

> 'But I must insist that $C(h, e)$ can be interpreted as degree of corroboration only if e is *a report on the severest tests we have been able to design*. It is this point that marks the difference between the attitude of the inductivist, or verificationist, and my own attitude.'

Thus Popper hardly seems seriously committed to an inductive interpretation of his degree of corroboration. Even if we do interpret Popper's degree of corroboration inductively, however, there still remains a considerable difference between this Popperian inductivism, and the Bayesian inductivism of Keynes and the Vienna Circle.

To begin with, Bayesian inductivists hold that the confirmation or corroboration function satisfies the ordinary axioms of probability. This is something which Popper denies.[3] Secondly, and this is the point which will principally concern us here, Keynes and the Vienna Circle believed that induction could be justified by appeal to a principle or principles of induction. Popper on the other hand has always repudiated such principles.

Let us briefly summarise Keynes' account of principles of induction as set out in his [1921] *A Treatise on Probability.* Keynes mentions first a familiar such principle—the so-called law of uniformity of nature. This, he claims, ([1921] Ch. 19.5 p. 252), involves: '. . . the irrelevance of mere position in time and space to generalisations which have no reference to particular positions in time and space. It is in respect of such position in time or space that "nature" is supposed "uniform".'

Keynes goes on to suggest that we may need something rather stronger than the law of uniformity of nature. This could be called the principle of atomic uniformity, and it is roughly described as follows ([1921] Ch. 21.5 pp. 276–7):

> 'The system of the material universe must consist . . . of bodies which we may term . . . *legal atoms,* such that each of them exercises its own separate, independent, and invariable effect, . . .'

Keynes introduces a third principle, that of the limitation of independent variety. However he remarks, [1921] Ch. 22.12 p 290: that 'The hypothesis of atomic uniformity . . .while not formally equivalent to the hypothesis of the limitation of independent variety, amounts to very much the same thing.' Keynes' view is that the natural sciences are founded on inductive inferences, and that these inductive inferences are only valid if the principles just mentioned hold in the world of nature. He concludes his *Treatise on Probability* by remarking optimistically, [1921] Ch. 33.10 p. 468:

> '. . . those peculiar characteristics of atomism and limited variety . . . appear more and more clearly as the ultimate result to which material science is tending . . .'

The development of the natural sciences in the 20th century does not bear out this claim of Keynes'. Einstein's general theory of relativity (1916) casts doubt on the Principle of Uniformity of Nature (as Keynes himself observes in a footnote [1921] Ch. 21.6 p. 276). The whole notion of 'legal atoms' has been called into question by the quantum mechanics of Heisenberg (1925) and Schrodinger (1926). Nor is there anything in the present scientific world picture to suggest the validity of a principle of limitation of independent variety. As material science is in fact tending to render dubious the Keynesian principles of induction, let us next examine Popper's attempt to dispense with such principles altogether.

---

[3] For further details see my [1988].

Popper's criticisms of principles of induction can be summarised in 3 groups.[4]

(1) Such principles of induction are designed to justify inductive inferences, but they fail in this task, because the question at once arises of how the principles themselves are justified. The suggestion that the principles are justified inductively by experience leads to a vicious circle, and there is little plausibility in claiming that the principles are known *a priori* to be valid.

(2) It is difficult to formulate the principles in a satisfactory way. If they are to be at all plausible, the formulation has to be vague, in which case the principles are useless. More precise formulations are, however, liable to exceptions.

(3) Such principles are neither used nor needed in actual scientific practice.

To substantiate (3), Popper has propounded his well-known methodology of conjectures and reputations. The scientist proposes theories which are subjected to severe tests. If a theory fails a test, an attempt is made to find a modification of the theory, or a new theory, which explains the test result satisfactorily. If a theory passes a number of severe tests and so becomes well-corroborated, it can be used as the basis for practical applications. The alleged principles of induction do not play any role in this process.

These arguments of Popper's seem to me convincing, and I cannot therefore agree with Pheby, who claims that Popper needs (p. 107) 'to employ something like the Principle of the Uniformity of Nature for scientific activity to be at all possible'. On the contrary it seems to me that one of Popper's most important contributions in this area was to have shown a way to avoid the difficult and fruitless task of trying to formulate and justify principles of induction. This point is not a purely theoretical one, for, as we shall see in the next section, it has implications for the assessment of Keynes' critique of econometrics.

4 KEYNES ON ECONOMETRICS

Keynes' writings on methodology and probability belong mainly to the early period which ends with the publication of his *Treatise on Probability* in 1921. In the late 30s, however, after the publication of *The General Theory*, Keynes made another, highly controversial, contribution to methodology. This was provoked by Tinbergen's pioneering development of econometrics. Keynes did not approve of this development, and was highly critical of what Tinbergen had done. Keynes' views on this subject are set out in his review of Tinbergen's book (Keynes [1939]), and in two important letters to R. F. Harrod on the methodology of economics (Keynes [1938]).

---

[4] For Popper's own treatment see, for example, his [1959] Ch. 1.1 pp. 27–30.

The contributors to the present volume differ in their assessment of Keynes' criticism of Tinbergen. Klant in Ch. 5 describes it as 'rather presumptuous', while Lawson in Ch. 7 and Pesaran and Smith in Ch. 8 endorse much of what Keynes says about econometrics. Keynes' criticisms of econometrics can be divided into those which deal with technical issues, and those of a more philosophical or methodological character. The technical criticisms are well-explained and illustrated by Pesaran and Smith in Ch. 8. There is no doubt that Keynes drew attention to some very real difficulties. Yet technical developments can very often overcome technical difficulties—at least in many cases, if not in all. Thus technical difficulties alone do not seem to justify Keynes' dismissal of econometrics as 'black magic' and 'statistical alchemy' (Keynes [1939]). Lawson's view is that Keynes' philosophical objections to econometrics are more important. As he says (p. 116):

> '... contributions have tended to concentrate on Keynes' comments on econometrics that deal with "technical issues" ... I wish to argue that the main thrust of Keynes' criticisms are of a more philosophical or methodological nature deriving from the theories and arguments which he developed in his *Treatise on Probability*.'

Now, as we saw in the previous section, Keynes thought that inductive inference was only justified when certain general inductive principles held good. Keynes' argument, Lawson claims, is that these inductive principles do not hold in the kind of cases with which econometrics deals. Statistical induction in such cases is therefore impossible. Indeed Keynes himself says, [1939] p. 316:

> 'For the main *prima facie* objection to the application of the method of multiple correlation to complex economic problems lies in the apparent lack of any adequate degree of uniformity in the environment.'

The problem, in short, seems to be that the law of uniformity of nature is not valid in such cases.

Pesaran and Smith agree with Lawson's reconstruction of Keynes' argument, and they summarise the position as follows (pp. 146–7):

> 'Some degree of homogeneity over space and time in the relationships considered is a prerequisite for statistical inference, just as the presumed constancy of natural laws, whether deterministic or probabilistic, is a prerequisite for physics...
>
> Given that there are not strong *a priori* reasons for believing economic relationships to be stable over time, and the fact that estimated equations are prone to structural change, one is forced to agree with Keynes that at a logical level econometric inference, like other forms of induction, is insupportable.'

I agree with the way in which Lawson, Pesaran and Smith reconstruct Keynes' argument, but not with their endorsement of this argument. Indeed, as I argued in the previous section, it seems to me that Popper has shown that scientific practice does not need the support of any inductive principles such as the law of uniformity of nature. Thus even if this law does not hold in the cases dealt with by econometrics, this is not an obstacle to our applying the usual procedures of science. Suppose, for example, that economic relationships do vary over time, why should we not form a conjecture as to how this variation takes place, and then test this conjecture against the data? This task may be a difficult one, but to say that a task is difficult is not to say that it is impossible.

Pesaran and Smith go on to say (p. 147): 'However, having given Keynes victory on both the technical and logical arguments, it still does not follow that econometrics is worthless.' Tinbergen advocated econometrics both for its practical usefulness in decision making and policy formation, and for the purpose of testing economic theories. Pesaran and Smith argue that econometrics has indeed (p. 148) a 'practical value to policy makers', but they regard 'the search for "truth" through testing economic theories' as an 'unattainable ideal'.

On this issue I side with Tinbergen against Pesaran and Smith, whose position seems to me to concede too much to instrumentalism. I cannot see why economic science should resign itself to being only a practical tool for the policy maker, and would rather advocate a search for truth about economic reality through the method of conjectures and refutations. Pesaran and Smith do, however, have an argument against the possibility of testing economic theories using econometrics, and this is what we must next consider. They state the argument as follows (p. 145)[5]:

> 'The econometrician . . . has the task of intermediate specification in order to obtain an equation which can be subjected to statistical analysis. This involves making a set of auxiliary assumptions, about other relevant factors, functional form, dynamic specification, etc., . . .
>
> A consequence of this procedure is that one cannot know whether the results of the statistical analyses reflect inferentially on the economic theory or on the auxiliary assumptions. It may be useful to reject a specific formulation or hypothesis, but the significance of this for the more general underlying economic theory will always be ambiguous.'

Philosophies of Science will immediately recognise this as the Duhem-Quine Thesis, cf. Duhem ([1914] and Quine [1951]). This thesis may be stated (in more Duhemian form) as follows. Suppose we are testing some high-level theory T. In order to derive an observation statement O, it may be necessary to add to T a set of auxiliary assumptions A say. So the logical situation is *not* $T \vdash O$, *but* $T \,\&\, A \vdash O$. Thus if experiment or observation establishes $\neg O$, we cannot conclude $\neg T$, but only either $\neg T$ or $\neg A$. In order to eliminate the anomaly, the scientist has to alter T or A, but he or she cannot know definitely which should be changed. It is a matter, as Duhem says, which needs to be decided by 'the good sense' (*le bon sens*) of the scientist.

---

[5] A similar argument is given by Klant on p. 94. Klant concludes, in agreement with Pesaran and Smith, (p. 96): 'I think that when Keynes rejected econometrics as a method of testing he was basically right, . . .'.

Now the point I would like to make is that the Duhem-Quine thesis applies just as much to theoretical physics as it does to economics and econometrics. Indeed Duhem first formulated the point for theoretical physics, and the standard example of the Duhem-Quine thesis is taken from mechanics as applied in astronomy. This example is the Discovery of Neptune in 1846 by Adams and Leverrier.

From Newton's theory T + auxiliary hypotheses, astronomers were able to calculate the theoretical orbit of Uranus (the most distant planet then known). This theoretical orbit did not agree with the observed orbit. This meant that either T or one of the auxiliary hypotheses was false. Adams & Leverrier conjectured that the auxiliary hypothesis concerning the number of planets was in error. They postulated a new planet Neptune beyond Uranus and calculated the mass and position it would have to have to cause the observed perturbations in Uranus' orbit. Neptune was duly observed on 23rd September 1846 52′ away from the predicted position.

This part of the story is quite well-known, but there were some subsequent events which are not without interest. Another difficulty which occupied astronomers at the time concerned the anomalous motion of the perihelion of Mercury, which advanced 38 seconds per annum faster than it should have done according to standard theory. Leverrier tried the same approach which had proved successful in the case of the Uranus anomaly. He postulated a planet Vulcan nearer to the Sun than Mercury, and with a mass, orbit etc. which would explain the advance in Mercury's perihelion. However no such planet could be found.

It was not till 1916, when Einstein replaced Newton's theory T by his own theory T′ of General Relativity that the anomaly in the motion of Mercury's perihelion was at last explained. We see that, although the Uranus anomaly and the Mercury anomaly were *prima facie* very similar, success was obtained in one case by altering an auxiliary hypothesis, and in the other by altering the main theory. All this goes to show that there is no magic formula for success in scientific research, and perhaps the scientist needs luck as well as Duhemian good sense.

The Duhem-Quine thesis certainly applies in Physics, but we should not therefore conclude that it is wrong to try to test physical theories. The situation seems to me similar in economics. I agree with Pesaran and Smith that the Duhem-Quine thesis applies in economics, but I disagree with their conclusion that 'testing economic theory' therefore becomes an 'unattainable ideal'. What the Duhem-Quine thesis shows, where it applies in either physics or economics, is that, in the face of an anomaly, we can never be sure whether to alter the underlying theory or some auxiliary hypothesis. Yet we do not know something: namely that at least one of our assumptions must be altered in order to restore the agreement between theory and observation. This knowledge provides a stimulus for research, and for the develoment of new and better views about the nature of reality under study. My conclusion is that Tinbergen's plan to develop econometric methods for testing economic theories is very far from being wrong in principle, and indeed is a very admirable and commendable project, even though there may be, in some specific cases, great technical difficulties in carrying it out.

DONALD GILLIES
King's College London

REFERENCES

DUHEM, P. [1914]: *The Aim and Structure of Physical Theory* Ch. VI Physical Theory and Experiment, and Ch. VII The Choice of Hypotheses. Atheneum.

GILLIES, D. A. [1973]: *An Objective Theory of Probability*, Methuen.

GILLIES, D. A. [1988]: 'Probability and Induction', Ch. 9 of *Encyclopedia of Philosophy*, ed. G. H. R. Parkinson. Croom Helm.

HOWSON, C. [1984]: 'Bayesianism and Support by Novel Facts', *British Journal for the Philosophy of Science*, 35, pp. 245–251.

JEFFREYS, H. [1939]: Theory of Probability. Oxford. 3rd Edition. 1961.

KEYNES, J. M. [1921]: *A Treatise on Probability* Vol. VIII of the Collected Writings of John Maynard Keynes. Macmillan. 1973.

KEYNES, J. M. [1933]: *Essays in Biography* Vol. X of the Collected Writings of John Maynard Keynes. Macmillan. 1972.

KEYNES, J. M. [1936]: *The General Theory of Employment Interest and Money* Vol. VII of the Collected Writings of John Maynard Keynes. Macmillan. 1973.

KEYNES, J. M. [1937]: 'The General Theory of Employment' in Vol. XIV of the Collected Writings of John Maynard Keynes. Macmillan. 1973. pp. 109–123.

KEYNES, J. M. [1938]: 'Correspondence with R. F. Harrod on the Methodology of Economics' in Vol. XIV of the Collected Writings of John Maynard Keynes. Macmillan. 1973. pp. 295–301.

KEYNES, J. M. [1939]: 'Professor Tinbergen's Method' in Vol. XIV of the Collected Writings of John Maynard Keynes. Macmillan. 1973. pp. 306–320.

LAWSON, T. [1985]: 'The Context of Prediction (and the Paradox of Confirmation)', *British Journal for the Philosophy of Science*, 36, pp. 393–407.

MUSGRAVE, A. [1974]. 'Logical versus Historical Theories of Confirmation', *British Journal for the Philosophy of Science*, 25, pp. 1–23.

POPPER, K. R. [1959]: *The Logic of Scientific Discovery* English translation. Hutchinson. Sixth Impression. 1972.

QUINE, W. V. O. [1951]: *From a Logical Point of View* Ch. II. Two Dogmas of Empiricism. Harper Torchbook.

RAMSEY, F. P. [1926]: 'Truth and Probability' in *The Foundations of Mathematics* Routledge & Kegan Paul. 4th Impression. 1965. pp. 156–198.

Brit. J. Phil. Sci. 39 (1988), 131–140 *Printed in Great Britain*

REVIEW

POLLOCK, JOHN [1986]: *Contemporary Theories of Knowledge.*

Rowman & Littlefield. *xii*+208 pp.

PAUL K. MOSER
Loyola University of Chicago

This book makes important original contributions to the theory of justified empirical belief. It argues in chapters 1–5 for a theory of epistemic empirical justification that improves on the others in circulation; and it defends in its appendix a notion of objective justification that provides a surprisingly straightforward diagnosis and resolution of Gettier-style counterexamples to the justified-true-belief analysis of knowledge. Due to space limitations, I cannot consider the appendix on the Gettier problem, but its approach seems to me basically correct and persuasive. (In fact, my own analysis of propositional knowledge in [1987b] has definite affinities to Pollock's.)

The book begins by outlining two major problems facing a theory of knowledge: the problem of explaining why skepticism about knowledge is wrong, and the closely related problem of explaining how it is possible to acquire various kinds of knowledge and justification. The latter problem leads Pollock to classify theories that explain how knowledge is possible as being either *doxastic* or *nondoxastic.* Doxastic theories endorse *the doxastic assumption,* according to which 'if one holds precisely the same beliefs in two possible circumstances, then no matter how these circumstances differ with respect to things other than what one believes, there will be no difference in what beliefs are justified under those circumstances' (p. 19). Nondoxastic theories, as would be expected, reject this doxastic assumption.

On Pollock's taxonomy a doxastic epistemological theory is either a foundations theory or a coherence theory. A foundations theory about empirical beliefs is committed to the view that every justified empirical belief either is an *epistemologically basic belief* or is justified ultimately on the basis of such a basic belief. Pollock adds:

> The basic beliefs themselves are not supposed to stand in need of justification. They are in some sense 'self-justifying'. One is automatically justified in such a belief merely by virtue of having it. (p. 20; cf. pp. 20, 28, 29, 35, 48, 56)

But this characterization of basic beliefs is too narrow if it intends to capture the historically prominent versions of foundationalism. And apparently it does intend to do so, since Pollock cites C. I. Lewis and others as proponents of a foundations theory in the sense specified (p. 26). However, early 20th-century foundationalists such as Lewis did not take foundational empirical beliefs to be 'self-justifying' or 'automatically justified'; rather, they located the justification for such beliefs in *non*-belief sensory and perceptual states, or at least in the non-judgmental content of such states, which Lewis and others called 'the given'. (For some of the details on Lewis's account, see Moser [1985], ch. 5, and [1987a].) Thus, a better way to characterize a basic belief is to say that it is justified, but is not justified by another belief or proposition. Given this characterization, a foundations theory need not be a doxastic theory. However, to avoid verbal disputes, let us grant Pollock his definition of foundationalism, while recognizing that it fails to capture some historically prominent versions.

A coherence theory contrasts with foundationalism in Pollock's sense, but also qualifies as a doxastic theory. The defining trait of a coherence theory is that it makes the justification of a person's belief a function solely of the way his belief 'coheres' with the other he holds. Pollock distinguishes between negative and positive coherence theories (p. 72). According to negative coherence theories, all beliefs are *prima facie* justified in the sense that one's holding a belief is sufficient for its being justified unless one has a reason not to hold it. Pollock attributes such a view to Harman [1986], although it seems questionable whether Harman really qualifies here. One consideration is that Harman's recent work on reasons is intended to answer questions about reasons for *belief-revision*, which Harman (p. 29) contrasts with questions about justification. In any case, it is noteworthy that Pollock himself defended a negative coherence theory, as an account of subjective justification, in his [1979]. (See chapter 3 of my [1985] for criticisms.) Positive coherence theories, in contrast, require positive support for the justification of a belief; specifically, they demand that a justified belief be supported by reasons that involve coherence relations to other beliefs in one's possession. These coherence relations are characterized in various ways by coherentists, but it is common to regard entailment and explanatory relations between beliefs as paradigm coherence relations. Also, it is now common for coherentists to endorse *holistic* (as opposed to linear) positive coherentism, according to which justification is a function not of coherence relations between individual beliefs, but of coherence relations between a belief and the set of all one's beliefs. (See, for example, Bonjour [1985] and Lehrer [1986].)

Pollock claims (p. 19) that the category of doxastic theories is 'exhausted' by foundations theories and coherence theories. But this is questionable, if 'exhausted' means 'logically exhausted'. Consider the view suggested by Wittgenstein [1969] and Kuhn [1970] that at the foundation of justified belief lies belief that is *un*justified. On this view, unjustified basic beliefs have at least two features in common with the basic beliefs of foundationalism: they are not justified by other beliefs, and they are the basis for all inferentially justified beliefs. However, unjustified basic beliefs also have a feature that distinguishes them from the basic beliefs of foundationalism: they are not themselves actually justified. (Recall here Wittgenstein's remark ([1969], par. 253) that 'at the foundation of well-founded belief lies belief that is not founded'.) I have argued elsewhere (in chapter 2 of [1985]) that such a Wittgensteinian view is unacceptable in its account of inferential justification. But the point now is that such a theory, when committed to the doxastic assumption, illustrates that foundations theories and coherence theories apparently do not logically exhaust the category of doxastic theories. However, I grant that Pollock can justifiably overlook the present view on the ground that it is initially uncompelling and seriously incomplete.

Nondoxastic theories of knowledge also occupy Pollock's conceptual geography. Naturally, all such theories reject the doxastic assumption, but they divide over the issue of what nondoxastic considerations are relevant to justified belief. Here Pollock distinguishes *internalist* from *externalist* nondoxastic theories (pp. 21–24). An internalist nondoxastic theory makes justification a function solely of one's internal states, including internal non-belief states, e.g., perceptual states. An externalist nondoxastic theory, in contrast, makes justification a function of more than just the internal states of a person. Here Pollock distinguishes *probabilism*, according to which the justification of a belief is a function of its probability of being true, and *process reliabilism*, according to which the justification of a belief is a function of the reliability, or truth-conduciveness, of the cognitive process that produced it. Clearly, since beliefs are internal states, every externalist theory is a nondoxastic theory. But not every internalist theory is a nondoxastic theory. In fact, foundations theories and coherence theories, as defined by Pollock, are paradigm versions of internalist doxastic theories.

Internal states, on Pollock's view, are one's states to which one can have 'direct access' in the sense that one can have access to them without first having beliefs or making judgments about them (pp. 22, 93, 133, 134, 136). Such access is nondoxastic inasmuch as it does not require prior belief about the states in question. But it is not clear what exactly such access involves. By way of example, Pollock says that when one learns to ride a bicycle, the 'automatic processing system' in one's brain has access to balance sensations emanating from the inner ear (p. 133). And he adds (p. 136) that 'a directly accessible property or relation is one to which our automatic processing system has access without our having beliefs about what things have the property or stand in the relation'. But some positive explanation of such access seems to be needed. Is such access purely nonconceptual? Or does it involve the forming of a belief or some more primitive form of conceptualization about the state, property, or relation in question? If so, does such access, being conceptual and belief-like, stand in need of justification, and what sort of justification is relevant here? On the other hand, if the access is not belief-like, how are we to understand *non*conceptual access? Since Pollock does not pursue such issues, his account is rather vague here.

None the less, we now can get a grip on Pollock's overall argument. It is simply that since all doxastic theories, specifically all foundations and coherence theories, and all externalist theories are false, we need to look to a version of nondoxastic internalism for a viable theory of justification. Let us consider the main steps of this argument and the sort of internalism recommended by it.

Pollock rightly observes that foundations theories do not require that a proposition be incorrigible to be epistemologically basic (p. 35). Also, he correctly emphasizes that foundations theories do not require logically conclusive reasons in the justification of nonbasic beliefs by basic beliefs. Nonconclusive reasons are defeasible in the sense that their support for a proposition can be defeated by considerations of a certain sort; in short, they are reasons for which there can be defeaters. Pollock calls such reasons *prima facie* reasons, and notes (p. 38) that they are susceptible to two kinds of defeaters: *rebutting* defeaters that provide one with a reason to believe the denial of the proposition otherwise supported by a *prima facie* reason, and *undercutting* defeaters that provide one with a reason to deny that the *prima facie* reason would not be true unless the otherwise supported proposition were true. Aside from familiar concerns about the subjunctive conditional here, Pollock's treatment of the role of *prima facie* reasons in relation to foundationalism is clear and enlightening. Yet Pollock finds that all foundations theories are eliminated by this argument:

> All beliefs (even beliefs about appearances) can be held for bad reasons. Depending upon what we say about what counts as a reason for thinking we should not hold a belief, this either has the consequence that no belief is prima facie justified or that all beliefs are prima facie justified. In either case, foundationalism fails, because foundationalism requires there to be a privileged subclass of beliefs that are prima facie justified and on the basis of which all other beliefs are justified (p. 65).

Given Pollock's narrow definition of foundationalism, this simple argument seems to refute all foundations theories. At least proponents of such theories now have a lot of explaining to do.

Pollock's major argument against coherence theories, like his main objection to foundationalism, takes issue with the doxastic assumption on which they rest (pp. 87–90). The argument is roughly as follows:

> 1. Given the doxastic assumption, facts about one's perceptual situation, such as how one is appeared to, are relevant to the epistemic status of one's perceptual beliefs about physical objects only if one believes those facts.
> 2. In normal cases one has no beliefs about how one is appeared to, and the beliefs one has prior to acquiring a perceptual belief about a physical object are not

sufficient to uniquely determine what perceptual belief one should acquire; for perception is not inference from antecedent beliefs.
3. So, in normal cases, according to a doxastic theory such as a coherence theory, perceptual beliefs about physical objects are automatically epistemically permissible, i.e., prima facie justified.
4. But even in a case where one has no beliefs about how one is appeared to, perceptual beliefs about physical objects can be held for bad reasons, in which case they are not prima facie justified.
5. Hence, the doxastic assumption is false, since it implies 3, which conflicts with the true proposition 4.

Pollock concludes further that coherence theories also are false, on the ground that they require the false doxastic assumption. The general thrust of this argument seems decisive. Put positively, it is that the justification of one's beliefs, particularly one's perceptual beliefs, depends on more that just what one actually believes. Nondoxastic experiential considerations, such as how one is appeared to, are crucially relevant to the justification of beliefs. It is this point that underlies the traditional isolation objection to coherentism, according to which such a view cuts justification off from one's perceptual experience (cf. Pollock [1974], pp. 27–28; Moser [1985], Ch. 3).

Pollock's final foe is externalism, including probabilism and process reliabilism. His particular criticisms of the latter two views are detailed, informative, and forceful. The major objection to probabilism is that there seems to be no appropriate kind of probability for use in a probabilist theory of knowledge (p. 113). Such a theory will be conceptually circular if it is formulated with the notion of *epistemic* probability, and each of the other prominent notions of probability (i.e., subjective, physical, and propensity notions) faces various special problems noted in chapter 4. The major objection to reliabilism is that there seems to be no plausible way to specify the set of circumstances relative to which the reliability of a cognitive process is justification-conferring (pp. 118–20). Cognitive processes are reliable under some circumstances, but unreliable under others. But how specific are we to make the circumstances where a reliable cognitive process is justification-conferring? If we include *everything* in the relevant circumstances, we have to consider the actual truth value of the belief produced, and thus the relevant indefinite probability will be either 1 or 0 depending on that truth value; but surely justification does not require truth. Pollock shows that the reliabilist seems not to have a non-arbitrary way to specify the circumstances in question.

Pollock's argument against externalism in general is more unified; in broad outline it is:

1. A belief is justified if and only if it is epistemically permissible (p. 124).
2. A belief is epistemically permissible if and only if it is held in conformance with correct epistemic norms (p. 125).

3. Correct epistemic norms must be 'reason-guiding' in the sense that they can guide us in our reasoning and belief-formation processes (p. 125).
4. Given externalism, the justification of a belief is a function in part of external considerations, and thus on that view external considerations determine in part whether a belief is held in conformance with correct epistemic norms. *Belief externalism* holds that the content of correct epistemic norms involves external considerations, whereas *norm externalism* states only that external considerations are necessary in the selection of correct epistemic norms (pp. 125–26, 138).
5. Belief-externalist theories are false, because their epistemic norms cannot be internalized in the way required for them to be reason-guiding; specifically, their norms do not involve only directly accessible properties, which can be accessed without the formation of prior beliefs (p. 136).
6. Norm-externalist theories are false, because external considerations, such as the objective reliability of a cognitive process, cannot bear on the reason-guiding justification of our beliefs; they are relevant only if they are internalized in the sense that we are apprised of them (p. 141).

Given 1–6, and the above argument against doxastic theories, Pollock is faced with the constructive task of explaining and defending a version of nondoxastic internalism as an alternative to foundationalism, coherentism and externalism.

Pollock's theory, called *direct realism*, is familiar from his [1974] and [1979]. Its basic assumption is that 'perceptual states can license perceptual judgments about physical objects directly and without mediation by beliefs about the perceptual states' (p. 91; cf. p. 175). On this view epistemic norms can appeal directly to one's perceptual states without referring to beliefs about perceptual states. An obvious advantage of this view, in contrast to traditional foundationalism, is that it fits with the fact that very few people actually form beliefs about their subjective perceptual states. Essential to direct realism are nondoxastic *prima facie* reasons such as the following:

> *X*'s looking red to *S* is a *prima facie* reason for *S* to believe that *X* is red.

On this view, perceptual states themselves convey *prima facie* justification. Further, the correct epistemic norms license inferences directly from perceptual states to physical-object beliefs, in cases where there are no defeaters. Such direct realism, as Pollock shows, is a clear improvement over the sort of foundationalism, coherentism, and externalism described above.

We can probe the basis of direct realism by asking how Pollock avoids challenges from epistemological relativism and skepticism. The relativist holds that (a) different people may have different epistemic norms that conflict inasmuch as they lead to different assessments of the justification of one and the same belief, and that (b) there is no non-arbitrary way to choose between such conflicting norms. The skeptic holds that we are not even *prima facie* justified in holding physical-object beliefs on the basis of our perceptual states. Pollock's reply to both challengers comes from his epistemological theory of concepts, which states that a concept is individuated by its role in reasoning. The basic idea is explained as follows:

> What makes a concept the concept that it is is the way we can use it in reasoning, and that is described by saying how it enters into various kinds of reasons, both conclusive and prima facie. Let us take the *conceptual role* of a concept to consist of (1) the reasons (conclusive or prima facie) for thinking that something exemplifies it or exemplifies its negation, and (2) the conclusions we can justifiably draw (conclusively or prima facie) from the fact that something exemplifies the concept or exemplifies the negation of the concept. My proposal is that concepts are individuated by their conceptual roles. The essence of a concept is to have the conceptual role that it does (p. 147; cf. Pollock [1974], ch. 1).

Given this theory of concepts, Pollock rejects relativism on the ground that if two people reason in accordance with different epistemic norms, they are actually employing different concepts (p. 148; cf. Pollock [1987], p. 84). The epistemic norms one uses in reasoning describe the conceptual roles of one's concepts and thus determine what concepts one uses. So, it is impossible for people to use different epistemic norms with regard to the same concepts. On this basis Pollock infers that epistemological relativism is logically false.

Skepticism meets a similar fate. On Pollock's view, it is a 'necessary truth' that our perceptual experiences provide us with justifying prima facie reasons for physical-object beliefs; and this necessary truth is due to the 'necessary features' of our concepts (i.e., their conceptual roles in reasoning) that make them the concepts they are (p. 178). Our holding physical-object beliefs on the basis of perceptual experiences is justified as a matter of logical necessity, according to Pollock, because that just is to ascribe physical-object concepts under their justification conditions, and those justification conditions are constitutive of those concepts (cf. Pollock [1967], p. 57). Further, Pollock finds it pointless to ask why our justification conditions are the relevant ones, since those conditions are constitutive of our concept of a physical object, and thus to introduce different justification conditions would be to introduce a different concept. In a similar vein, Pollock claims that 'our epistemic norms are constitutive of the concepts we have and hence it is a necessary truth that our actual epistemic norms are correct' (p. 168). Such considerations are not intended to prove logically that we have *knowledge* of physical objects, but they do purport to show that we are justified in believing that certain things are true of the physical world.

Pollock's theory of concepts raises many important issues, only one of which will be considered here with regard to skepticism. The basic question is whether our concept of a physical object is actually constituted, or essentially determined, by its conceptual role, i.e., by our justification conditions as specified by direct realism. Could it be, in other words, that our concept of a physical object outstrips our justification conditions for it inasmuch as those conditions fail to bear on a central feature of that concept? Suppose the skeptic proposes that (a) our concept of a physical object essentially involves the subjunctive conditional that an object falling under that concept would exist even when unperceived, and that (b) nothing in one's merely being appeared to by a physical object (i.e., an apparent physical object) supports, in *any* way, such a subjunctive conditional. From (a) and (b) the skeptic might infer that one's merely being appeared to by a physical object does not justify one's ascribing the concept of a physical object. The direct realist might reply that the conceptual role of the concept of a physical object consists of (1) the justifying reasons for holding that something exemplifies it or its negation and (2) the conclusions we can justifiably draw from the fact that something exemplifies the concept or its negation. But constituent (1) is not decisive here, since the skeptic is questioning whether the conditions specified by direct realism are actually correct justification conditions for the ascription of the concept of a physical object. And condition (2) is even less helpful, since the skeptic, being a radical internalist, questions whether any conclusions can be justifiably drawn from the mere fact that something exemplifies the concept of a physical object or its negation. Further, even if he grants that *if* one is justified in ascribing the concept of a physical object, then one is also justified in ascribing the concept of something that would exist when unperceived, he still questions whether direct realism correctly states the justification conditions for those concepts. For, again, he does not find in direct realism any support for the relevant subjunctive conditional. Consequently, the skeptic apparently would be unmoved by the appeal of direct realism to the epistemological theory of concepts. (cf. Moser [1988b].)

Further, we obviously cannot challenge the skeptic by revising constituents (1) and (2) to concern (1a) what we *believe* to be the justification conditions for the concept of a physical object and (2a) what we *believe* to be justifiable conclusions from something's exemplifying the concept of a physical object or its negation. For mere belief, as noted above, can be epistemically irrational, in which case it is hardly an effective challenge to the skeptic. Moreover, we evidently cannot introduce considerations of justified belief here, since the nature of correct justification conditions is now under dispute by the skeptic. Thus, it seems to be a wide open question whether Pollock can justifiably assume that 'we know from the start that the skeptic is wrong' (p. 179; cf. pp. 6, 15). My own view is that the only way to show why the skeptic is wrong is to appeal, not simply to justification conditions, but to explanatory considerations, such as that one's perceptual experiences are ordinarily *best explained* by physical-object beliefs. (See Moser [1988a].) Further, it is not obvious that the introduction of considerations of best available explanation into our (i.e., direct realist) justification conditions for ascribing the concept of a physical object would thereby alter that concept. (This point bears on Pollock's criticism of relativism, but cannot be developed here.)

I find that there are two other features of Pollock's direct realism that raise important questions. First, direct realism apparently requires not only nondoxastic *prima facie* reasons, but also *nonconceptual prima facie* reasons. For if one's epistemically basic perceptual states, such as *X*'s looking red to one, are essentially conceptual items whereby one ascribes a concept, they themselves will be in need of justification from something else, and thus will not really be basic after all. In order to avoid the threat of a regress here, the direct realist apparently is well-advised to endorse nonconceptual perceptual states. But these big questions will then arise: Do such nonconceptual perceptual states really exist? If so, what is the evidence for their existence? Further, what distinguishes them from nonperceptual states of mere sensory stimulation? And how can a nonconceptual perceptual state play any role in reason-guiding justification? Direct realism must provide full answers to such basic questions. Second, Pollock's direct realism states that our mere beliefs about our nondoxastic states take precedence over our nondoxastic states in determining what we are justified in believing. Thus, 'if *X* looks red to me, but I believe that *X* does not look red to me, then I am not justified in taking *X* to be red on the basis of the perceptual state' (p. 177). But this seems implausible. Why should a mere, *un*justified belief take epistemic precedence over a justifying perceptual state? Beliefs can be held for bad reasons and when they are, they should not take epistemic precedence over perceptual states. It seems, then, that genuine defeaters must themselves be epistemically supported; otherwise, defeaters can be created arbitrarily.

Despite my unanswered questions about direct realism, I recommend Pollock's book as a first-rate treatment of justified empirical belief. The book can be read with great profit not only by advanced philosophy students, but by all professional philosophers. In fact, as far as epistemology books go, *Contemporary Theories of Knowledge* is required reading.

REFERENCES

Bonjour, L. [1985]: *The Structure of Empirical Knowledge.* Harvard University Press.

Harman, G. [1986]: *Change In View.* MIT Press.

Kuhn, T. S. [1970]: *The Structure of Scientific Revolutions, 2d ed.* University of Chicago Press.

Lehrer, K. [1986]: 'The Coherence Theory of Knowledge'. *Philosophical Topics* **14**, pp. 5–25.

Moser, P. K. [1985]: *Empirical Justification.* D. Reidel Publishing Company.

Moser, P. K. [1987a]: 'Foundationalism, the Given, and C. I. Lewis', *History of Philosophy Quarterly* **5**.

Moser, P. K. [1987b]: 'Propositional Knowledge', *Philosophical Studies* **52**, pp. 91–114.

Moser, P. K. [1988a]: 'The Foundations of Epistemological Probability', *Erkenntnis* **29**.

Moser, P. K. [1988b]: 'Against Semantic Justificationism', *Philosophical Papers* 17.

British Society for the Philosophy of Science

Annual Conference 1988
University of Hull: September 23–25 1988

PROFESSOR IAN HACKING (University of Toronto):
Title to be announced

DR. JEREMY BUTTERFIELD (University of Cambridge):
Peaceful Coexistence: How's Chances?

DR. HARVEY BROWN: *A Reply*

PROFESSOR SUSAN HAACK (University of Warwick):
What is 'the Problem of the Empirical Basis' and does Johnny Wideawake solve it?

PROFESSOR JOHN WATKINS (L.S.E.): *A Reply*

PROFESSOR ANTHONY O'HEAR (University of Bradford):
Evolution and Knowledge

MR. MICHAEL SMITHURST (University of Southampton):
The Elusiveness of Human Nature

PROFESSOR JOHN PHEBY (Birmingham Polytechnic):
The Relevance of Laudan for Economists

DR. SHEILA DOW (University of Stirling):
Keynseian Thought—An Alternative to Dualism

Papers by Research Students

Conference Organiser: Dr. A. J. Dale
Department of Philosophy,
University of Hull,
Hull, U.K.

CAMBRIDGE

From Paradox to Reality

Our Basic Concepts of the Physical World

FRITZ ROHRLICH

This book discusses, in clear non-technical language, the two major theories of modern physics, relativity and quantum mechanics. The author offers a basis for the discussion of many strange new concepts and of the scientific necessities that led to them.

227 pp. 1987 0 521 30749 X **£25.00 net**

Particles and Paradoxes

The Limits of Quantum Logic

PETER GIBBINS

The fierce debate about quantum theory, its challenges to the laws of classical logic and its suggestion that the physical world is illogical are reviewed in this valuable introduction to the philosophy of quantum mechanics.

181 pp. 0 521 33498 5 **Hard covers £25.00 net**
0 521 33691 0 **Paperback £8.95 net**

Niels Bohr's Philosophy of Physics

D. R. MURDOCH

Giving a clear and forthright exposition and analysis of Niels Bohr's ideas, this book explores the often subtle implications of Bohr's original writings and indicates their prime importance in the formulation of the Copenhagen interpretation of quantum physics.

'The breadth and clarity of Murdoch's analysis makes this book a worthy addition to Bohrian scholarship.' *New Scientist*

304 pp. 1987 0 521 33320 2 **£32.50 net**

Gassendi the Atomist

Advocate of History in an Age of Science

LYNN S. JOY

This study of the life and work of the French physicist and philosopher, Pierre Gassendi, describes how he revived the ancient Greek philosophy of Epicurean atomism and points to the important role of humanist scholarship in the development of science in the seventeenth century.

328 pp. 0 521 30142 4 **£27.50 net**

Ideas in Context

Philosophical Issues in Aristotle's Biology

Edited by **ALAN GOTTHELF** and **JAMES LENNOX**

The theme of this collection of essays is the value of Aristotle's biology for the study of his philosophy, and the aim is to introduce scholars and students to the best work done in this area. The contents are divided into four parts, each part with an introduction which places its essays in relation to each other and to the wider issues of the book as a whole.

480 pp. 0 521 32582 X **Hard covers £30.00 net**
0 521 31091 1 **Paperback £10.95 net**

For further details please write to Shelby Howe at the address below.

Cambridge University Press
The Edinburgh Building, Shaftesbury Road, Cambridge CB2 2RU.

Synthese

An International Journal for Epistemology, Methodology and Philosophy of Science

Editor-in-Chief

Jaakko Hintikka, *Dept. of Philosophy, Florida State University, Tallahassee, USA*
Synthese publishes articles in all the fields covered by the subtitle. These include the theory of knowledge; the general methodological problems of science, such as the problems of scientific discovery and scientific inference, those aspects of symbolic logic and of the foundations of mathematics which are relevant to the philosophy and methodology of science; and those facets of the history and sociology of science which are important for contemporary topical pursuits.

Subscription Information **ISSN** 0039–7857
1988, Volumes 74–77 (12 issues)
Institutional rate: Dfl. 904.00/US$428.00 incl. postage/handling
Private rate: Dfl. 344.00/US$144.00 incl. postage/handling
Single issues: Dfl. 45.00/US$21.50 incl. postage and handling

D. Reidel Publishing Company

Kluwer Academic Publishers

P.O. Box 989, 3300 AZ Dordrecht, The Netherlands
101 Philip Drive, Norwell, MA 02061, U.S.A.

History of Science

the only review of literature and research in the history of science, medicine and technology in its intellectual and social context

1962-1987

A special post-free offer to new subscribers
valid until 31 December 1988

Volumes 1-25 hardbound, US $630/£315

From Science History Publications Ltd, Halfpenny Furze,
Mill Lane, Chalfont St Giles, Bucks HP8 4NR, England

New from Oxford

Particles and Ideas

Bishop Berkeley's Corpuscularian Philosophy

Gabriel Moked

Demonstrates how Bishop Berkeley, contrary to popular opinion, developed a highly original brand of corpuscularian physics.

0 19 824990 X, 256 pages, Clarendon Press £27.50
June 1988

The Reality of Numbers

A Physicalist's Philosophy of Mathematics

John Bigelow

This book casts new light on mathematics through its consideration of metaphysical materialism and set theory.

0 19 824957 8, 208 pages, Clarendon Press £25.00
June 1988

Philosophical Foundations of Quantum Field Theory

Edited by **Harvey R. Brown and Rom Harré**

In this unique study of the nature and implications of quantum field theory the contributors explore its mathematical structure and work through its metaphysical and methodological implications.

0 19 824975 6, 208 pages, Clarendon Press £25.00
March 1988

The Description of Nature

Niels Bohr and the Philosophy of Quantum Physics

John Honner

Bohr was as original a philosopher as he was a physicist. This study explores several dimensions of his vision.

0 19 824976 4, 256 pages, Clarendon Press £25.00

The Metaphysics of Quantum Theory

Henry Krips

The interplay between non-relativistic quantum theory and metaphysics has generated radically opposed interpretations for quantum theory. This book outlines the contours of these debates and presents an interpretation which resolves most of the paradoxes.

0 19 824971 3, 272 pages, Clarendon Press £27.50

The Scientific Image

Bas. C. van Fraassen

Winner of the Lakatos Award for Philosophy of Science presented at the London School of Economics, February 1987.

'. . . makes a clear statement of what is involved in scientific realism, and it will become the classic current argument against that philosophical position.' *Nature*

Clarendon Library of Logic and Philosophy
0 19 824424 X, 248 pages £22.50
0 19 824427 4, paper covers £9.95

Incompleteness, Nonlocality, and Realism

A Prolegomenon to the Philosophy of Quantum Mechanics

Michael Redhead

Concentrating on research done during the last twenty years, the author focuses on three major issues: whether quantum mechanics is an incomplete theory, whether it is non-local, and whether it can be interpreted realistically.

'Redhead's book is a useful update.' *New Scientist*

0 19 824937 3, 200 pages, Clarendon Press £25.00

Oxford University Press

REVUE INTERNATIONALE DE PHILOSOPHIE

Editor: Michel MEYER
143, av. A. Buyl, 1050 Brussels, Belgium

Each number is devoted to a particular movement, a particular philosopher, or a particular problem.

We publish 3 issues annually. Articles are written in English, French, German, or Italian.

Coming numbers

Locke
Schopenhauer
Kant and Practical Reason
Artificial Intelligence
Heidegger
Wittgenstein

Our last issue was devoted to *Philosophical Aspects of Literary Criticism*

(articles by M. Bloomfield, M. Brinker, J. M. Ellis, M. Frank, G. Hartmann, W. Iser and M. Meyer, D. E. Wellbery and E. M. Zemach)

Philosophical Studies

An International Journal for Philosophy in the Analytic Tradition

Editor-in-Chief
John Pollock, *Dept. of Philosophy, University of Arizona, Tucson, USA*

Philosophical Studies is devoted to the publication of analytical contributions, particularly (but not exclusively) in epistemology, philosophical logic, the philosophy of language, and ethics. The principal aim is to publish articles that are models of clarity and precision in dealing with some significant philosophical issues. Articles in the journal are intelligible to philosophers whose expertise lies outside the subject matter of the article.

Subscription Information **ISSN** 0031–8116
1988, Volume 53–54 (6 issues)
Institutional rate: Dfl. 422.00/US$200.00 incl. postage/handling
Private rate: Dfl. 164.00/US$69.00 incl. postage/handling

Private subscriptions should be sent direct to the publishers

D. Reidel Publishing Company

Kluwer Academic Publishers

P.O. Box 989, 3300 AZ Dordrecht, The Netherlands
101 Philip Drive, Norwell, MA 02061, U.S.A.

History of Science is published quarterly in issues of about 112 pages.

Volume 26 (1988) will include

- A. C. CROMBIE *Designed in the Mind: Western Visions of Science, Nature and Humankind*
- GEORGE WEISZ *The Self-made Mandarin: The* Eloges *of the French Academy of Medicine, 1824-47*
- L. A. CALLENDER *Gregor Mendel: An Apponent of Descent with Modification*
- J. C. D. BRAND *The Scientific Papers of C. P. Snow*
- N. ROLL-HANSEN *The Profession of Eugenics*
- JOSEPH CARON *Biology in the Life Sciences*
- KAREN PARSHALL *'The Art of Algebra'*
- F. W. P. DOUGHERTY *Correspondence of Lichtenberg*

The Annual Subscription is US $90.00 post-free in the Americas and Japan, £45.00 elsewhere ($44.00/£21.50 direct to private subscribers). Write to: Science History Publications Ltd, Halfpenny Furze, Mill Lane, Chalfont St Giles, Bucks HP8 4NR, England

the review of

metaphysics

a philosophical quarterly

MARCH 1988 | VOL. XLI, No. 3 | ISSUE No. 163 | $9.00

articles

STEPHEN WATSON	*Heidegger, Rationality, and the Critique of Judgment*
MICHAEL MURRAY	*Husserl and Heidegger: Contructing and Deconstructing Greek Philosophy*
JOHN CAPUTO	*Demythologizing Heidegger: Aletheia and the History of Being*
MICHAEL WILLIAMS	*Scepticism without Theory*

critical studies

DANIEL DAHLSTROM	*Heidegger's Last Word*
ALISTAIR MORRISON and Staff	*Summaries and Comments*

philosophical abstracts *announcements*

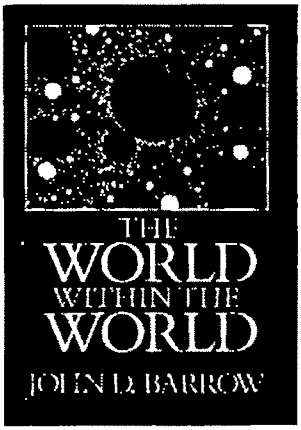
Individual Subscriptions $23.00 Institutional Subscription $35.00 Student/Retired Subscriptions $10.00

Jude P. Dougherty, Editor
The Catholic University of America, Washington, D.C. 20064

NEW FROM OXFORD

The World Within the World

John D. Barrow



Do there really exist laws of Nature?

Do they have an objective reality, or have we just invented them to make sense of our experience?

Why are they so well described by mathematics?

Is it even possible that there really are no laws of Nature—that they are an illusion?

How does our own existence limit what can be known about the Universe?

This book presents a wide-ranging interdisciplinary study of the evolving concept of laws of Nature. From the magical notions of primitive cultures to the latest ideas about chaos, black holes, inflation, and superstrings the author traces the gradual development of our concept of what laws of Nature are and how we come to know them. The scientific and mathematical topics discussed are of major importance to philosophers and scientists. This book gives them the opportunity to see these new ideas discussed in a serious but non-technical style within a historical context.

0 19 851979 6, 420 pp., illus., Clarendon Press,
May 1988 £20.00

New in Paperback

The Anthropic Cosmological Principle

John D. Barrow and Frank J. Tipler

This book addresses a central problem for both scientists and philosophers: the position and role of man in the universe.

'a remarkable book and a masterly exposition of . . . one of the most important developments to have taken place in physical science.' *Times Literary Supplement*

'an engaging book . . . will be much quoted, much debated, and much praised.' *Nature*

Oxford Paperbacks
0 19 282147 4, 640 pp., illus., April 1988 £9.95
0 19 851949 4, hardback, 1986 £27.50

OXFORD UNIVERSITY PRESS

Brit. J. Phil. Sci. 39 (1988), 141–160 *Printed in Great Britain*

Solving the Problem of Induction Using a Values-Based Epistemology

BRIAN ELLIS

---

1 *Introduction*
2 *The Empiricist Model*
3 *Improving the Empiricist Model*
4 *Induction and Epistemic Conservatism*
5 *Corroboration*
6 *Uniformity of Nature*
7 *Connectivity*
8 *Solving the Problem*

I INTRODUCTION

The problem of induction is to show that our scientific inductive practices are more or less rational. To do this, I shall argue, we need a theory of rationality in which rationality is defined in terms of optimal strategies for maximising epistemic value. The principal aim of this paper is to sketch such a theory, and to explain how it can be used both to justify our inductive practices, in so far as they are rational, and to explain them.

The new theory is one which makes certain epistemic values fundamental, and construes inductive rules as being just more or less useful strategies for promoting these values. It is thus a *values-based* epistemology. The theories sought by empiricists have usually been *rules-based*. For their aim has been to show that some particular inductive rule, e.g. the straight rule, is intrinsically rational, and hence that its use is rationally justified. However, for reasons given in Ellis [1965], it is not enough to show that some abstractly considered rule has intrinsically desirable properties. As Goodman has shown, what has to be justified is the way we actually use the rule. And this cannot be done without appeal to our epistemic values.[1]

---

[1] In Ellis [1965], I argued, much as Nelson Goodman did in Goodman [1954], for the importance of *theoretical conservatism* as an epistemic value, and claimed that our inductive rules must be applied in theoretically conservative ways if their use is to be rational. More recently, John Clendinnen [1982] and Wesley Salmon [1982] have argued that *simplicity* and *non-arbitrariness* are the main considerations in determining whether an inductive rule, such as the straight rule, has been rationally applied—thus giving these epistemic values a central role in their theory.

The hope of many philosophers is that the problem of induction can be solved by developing a naturalistic theory of rational competence adequate to explain and justify some of our inductive reasoning. I share this hope; but I think such a theory will have to be a values-based theory within which rules have only the status of epistemic strategies, or rules of thumb—much as rules do in act-utilitarianism. First, the inductive rules we use appear to have precisely this character. For there are no inductive rules which can be used indiscriminately, or where the subject matter is theoretically isolated. (See Ellis [1965]). Secondly, it has now been well established empirically that the inductive principles we find most intuitive are often not rational, and that the inductive practices we think are most rational are often not intuitive.[2] There is some reason to doubt whether the experimental findings warrant all of the conclusions that have been drawn about the shortcomings of intuitive reasoning (see Cohen [1986]). Even so, the prospects for developing an adequate epistemology which appeals to intuitively valid *principles* of inductive reasoning do not appear to be good.

Ultimately, the justification of our scientific inductive practices must depend somehow on our intuitive judgements of rationality, as Jonathan Cohen has successfully argued.[3] But the appeals to intuition we make when we evaluate inductive rules or strategies are generally to the rationality or otherwise of *particular inferences*, and so, presumably, to the perceived, epistemically valued, *properties* of these inferences. The appeal is not to any more intuitively evident *principles* which may serve as a basis for criticism. Therefore, if our theory is to reflect our practice, what we need is a values-based epistemology rather than a rules-based one. Compatibly with the particularism of our practice of epistemic evaluation, I suppose one could have a kind of Rossian epistemology (cf. Ross [1939]), in which inductive rules had the status of being *prima facie* rational. But I wish to put greater emphasis on the role of epistemic values in epistemology than Ross did on moral values in ethics.[4]

The nearest ethical parallel to my position in epistemology is ideal act-utilitarianism. In act-utilitarianism, the right act is that which maximises moral value in the long run. Somewhat analogously, I take the true belief to be that which would maximise epistemic value, if the matter were inquired into with perfect thoroughness by an ideally rational being (assuming no restriction of access to the kinds of information we are in principle capable of obtaining by observation or experiment). Given that the rightness of an act depends on its long-term consequences, judgements of right and wrong are normally defeasible. Consequently, we must distinguish between what it is morally right to do and what we are morally justified in doing. Something similar is the case with judgements of truth and falsity, because perfect thoroughness and ideal rationality are idealisations of the kinds of thoroughness and reasoning of which we are actually capable. Therefore, it is necessary to distinguish between truth and epistemic justification. The truth is what it is epistemically right to believe, not what we may be epistemically justified in believing.

---

[2] For a summary of the major findings on the shortcomings of intuitive reasoning, and an excellent discussion of them, see Nisbett and Ross [1980].

[3] See Cohen [1981] and [1986].

[4] Good discussions of the varieties of intuitionism in meta-epistemology are to be found in Alston [1978] and especially Dancy [1982].

This theory of truth is a pragmatic one. But it differs from other theories in this tradition by the nature of the epistemology on which it is founded. For there are no inductive rules in this epistemology which have more than the status of epistemic strategies. The truth, I want to say, is what it is right to believe in relationship to a specific system of epistemic values, and our rules of inference are just more or less good strategies for constructing the kinds of beliefs and belief systems we value most epistemically.[5]

The viability of this theory of truth clearly depends on the development of an adequate theory of rationality to define the concept of an ideally rational being. Therefore, the kind of theory we need is a model theory of the sort that is very common in science. This theory must set out to define an ideally rational being, and then, having done so, to explain how people actually think by contrast with how they would think if they were ideally rational.

The requirements on the model are:

1. The ideal of rationality contained in the model should appeal to the normal intellect as a rational ideal.
2. We should be able to use the model successfully to explain our inductive practices, i.e. our inductive practices should be more or less rational, as this is defined in the model; and in so far as they are not, we should be able to explain why.
3. The model should be biologically realistic, i.e. it should be explicable biologically why we should have the basic epistemic values we do.

A theory of rationality which had all of these features would provide an adequate solution to the problem of induction. For it would explain both why we think inductively as we do, and how and why rationally we ought to think inductively. I assume that, biologically, we are value-driven decision systems, and accordingly that the mechanism which shapes our belief systems is value-driven[6]. That is, I suppose that we have certain inbuilt values which together determine our belief preferences. Among these, I assume that there are certain rational, or epistemic, values which are distinguishable, although perhaps not sharply, from our personal and social values. Our *epistemic* values determine what we ought *rationally* to believe, as opposed to what is otherwise in our interests, or in the interests of society, or is pleasing or comforting to believe.

---

[5] This theory of truth, which I have elsewhere called an 'evaluative theory' (Ellis [1980]), distinguishes my position from Isaac Levi's. For in his development of a values-based epistemology, Levi supposes truth itself to be an epistemic value. (See Levi [1967] and [1980]). This is something I specifically deny. (See Section 3 below).

[6] The case for this has been argued by George Pugh. (See Pugh [1978].)

It is important to distinguish our epistemic values from the biological ends they serve. To flourish, people need reliable, economical, accurate and efficient belief systems which are well adapted to their environments, and applicable to most new situations which may arise. They also need to share many of the beliefs and attitudes of their tribe. Presumably, our epistemic values have been selected to make it possible for us to construct and maintain such belief systems. However, these biological needs are not our epistemic values, and it is not possible to justify our inductive practices with reference to them without circularity. For such ends refer beyond ourselves to our environments, and hence we must have knowledge of our environments if we are to use them for this purpose. Moreover, they may change or disappear as our circumstances change, but our epistemic values are not so externally dependent.

Our most basic epistemic values must be internally assessable things like consistency in our belief systems, beliefs that are empirically certified, well-corroborated beliefs, established beliefs and concepts, beliefs about what is universally the case, and theories establishing certain kinds of links between beliefs or systems of beliefs. Our biological needs have no doubt determined our epistemic values, but we must pursue them for their own sakes, without regard to this fact. It will be argued here that the values just listed are among those which motivate and explain our inductive practices. They are not all equally fundamental, and the list is certainly not complete. But they are all, I think, more or less basic human values which are independent of our social and personal values.

In pursuing knowledge and understanding, we seek to maximise the satisfaction of our epistemic values. And, to the extent that our inductive practices are effective means to this end, I claim that they are rational. For rationality is just the pursuit of epistemic value by appropriate means. Of course, many philosophers will want to define knowledge as knowledge of the truth, and to define truth in some way independently of our epistemic values. To such philosophers, I can offer no prospect of a solution to the problem of induction, because their position is essentially sceptical. But to those who think, as I do, that the truth is what it is right epistemically to believe, a naturalistic solution to the problem of induction may be possible. For the attempt to improve our knowledge and understanding in accordance with our epistemic values, and the quest for truth, are one and the same thing.

2 THE EMPIRICIST MODEL

Empiricists, like most other philosophers, consider *empirical certification* and *consistency* to be basic criteria for rationally evaluating beliefs and belief

---

[7] How consistency may be defined epistemologically is explained in Ellis [1979].

systems. I call these our *empiricist values*. According to many empiricists, these values are derivative. They are not innate, as I should say, but derive from the supposedly primary values of truth and avoidance of error.[7] Thus, consistency is to be valued, not for its own sake, but because it is a necessary condition for truth. Empirically certified (i.e. perceptually and introspectively acquired) beliefs are generally to be given epistemic priority, not because we have an inbuilt disposition to do so, as I would suppose, but because they are assumed to be known most directly to be true, and hence least prone to error.

According to many empiricists, truth and avoidance of error are the only genuinely primitive epistemic values. So that, on their view, if anything else has epistemic value, its value is derivative. The values of comprehensiveness, verisimilitude and validity, for example, are all said to be explained in terms of these fundamental values. Thus, a more comprehensive theory is generally to be preferred to a less comprehensive one, if neither has been falsified, not because it unifies and makes comprehensible a wider range of phenomena, but because, potentially at least, it contains more truth. As Popper would say, it says more. Verisimilitude, however it is to be measured, is to be valued, not because the criteria we use to gauge it refer to properties which have any intrinsic worth, but because it is supposed that theories possessing these properties are nearer to the truth. Likewise, valid arguments commend themselves for no reason other than that they are necessarily truth preserving, so that if we start with truths we shall never be led into error—although it is not explained why we should prefer to argue validly when we are dealing with ethical propositions, or subjective probability judgements, which, on many empiricist theories, are neither true nor false.

The problem of induction is one that arises within this empiricist framework. For inductive arguments are, by definition, not necessarily truth preserving, and hence invalid. The problem, then, is to explain why an ideally rational being, whose only primary concerns are with truth and avoidance of error, should accept such arguments. The conclusion reached by Hume, and never subsequently overturned, is that there is no rational justification for doing so. Consequently, it is now widely accepted by empiricists that our general beliefs about the world—those which do more than just summarise their known instances—are not rationally held. We hold them, it is said, not because we are rationally entitled to, but because we are creatures of habit, or have found them useful instruments for prediction, and have nothing better to put in their place, or because we have found them unifying, elegant or aesthetically pleasing—at any rate not for reasons which would, in themselves, justify us in believing them to be true.

Now a good theory always accounts for most of the facts in its domain. It may alter somewhat our perception of the facts, and hence our judgement about what a good theory in the area should do. But if a theory restricts the domain of the facts too much, or leaves unexplained most of what it was designed to explain, then this is reason to think that the theory is unsatisfactory. A planetary theory which explained only the motions of the outer planets would not be a good theory. It might be acceptable as a step in the direction of understanding planetary motion, but the research programme it defined, of accounting for what the theory left unexplained, would have to be fruitful for it to remain so.

The empiricist theory of rationality, which purports to explain what we may rationally believe, is not a good theory by this criterion. For it leaves most of what it set out to explain, viz., the rationality of science, to be explained in other ways, although no satisfactory explanations have been forthcoming. It can, perhaps, explain our preference for empirically certified beliefs, and certain general facts about the structure of our belief systems. But it does not explain why, rationally, we should accept what most of us would regard as scientifically established. On the contrary, it implies that most of our scientific beliefs are not rationally held, and if nevertheless we hold them, then an empiricist is committed to explaining this fact as due to the influence of non-rational forces, e.g. force of habit, aesthetic preference, or perceived instrumental value. The failure of empiricists to solve the problem of induction is, therefore, a powerful argument against empiricism.

3 IMPROVING THE EMPIRICIST MODEL

To construct a more adequate theory of rationality, we must replace the empiricist model by a better one. To do this, we must try to retain those features of the empiricist model which have explanatory power, or replace them by others capable of doing the same job. Specifically, the values of consistency and empirical certification must be retained, whether or not they are regarded as fundamental.

I think the first step towards improving the model is to reject the assumption that our primary epistemic values are truth and avoidance of error. For we need to break away from the traditional restraints, and there is reason to think that these are not in any case genuine epistemic values. If we reject them, however, we must somehow retain our empiricist values as features of the model, for otherwise we shall just be throwing away the baby with the bath water. My proposal for keeping the baby is that we should regard our empiricist values as primary ones, which are not to be justified in terms of other values, epistemic or otherwise. For, in this way, we can retain those features of the empiricist model which are explanatory without importing the metaphysical concept of truth into the foundations of our theory.

One reason I want to say that truth and avoidance of error are not fundamental epistemic values is that on any theory of truth, other than a subjectivist one, they are not internally assessable, and cannot therefore be the empistemic values which form and shape our belief systems from the start. It might well be biologically necessary for us to have beliefs which are mostly true, and so likely that our epistemic values are adapted to this end. But given any objectivist theory of truth, truth cannot be a fundamental epistemic value.

But I want to go further than this. In my view, truth and avoidance of error are not epistemic values at all. We use various criteria to evaluate beliefs as true or false, but we do not use truth or falsity as criteria for the evaluation of beliefs in any other respect. 'True' and 'false' are terms which function in epistemology like 'right' and 'wrong' do in ethics. They are what I have elsewhere called 'modes of evaluation'. (Ellis [1980]) Now moral rightness is not a moral value, but something which has to be casked in terms of moral values, which are things like justice, happiness, considerateness, honesty, and so on. If we want to understand what moral rightness is, we must understand the value system with respect to which our judgements of right and wrong are made. Likewise, I should say, truth is not an epistemic value, but a mode of evaluation for a system of epistemic values. And if we want to know what truth and falsity are, we must understand the system of epistemic values with respect to which our judgements of truth and falsity are made.

Having made this move, the question then becomes: What are our epistemic values, and how are they interrelated? To retain the virtues of empiricism, I assume that consistency and empirical certification must certainly be two of them, and that both have high priority. Indeed, I should suppose that the value of consistency always overrides. It is clear, however, that we need other values besides these to explain our belief preferences, and hence our judgements of truth, falsity, probability, and so on. For example, the assumption that the world did not exist before people existed is compatible with our empiricist values, as is the assumption that the sun will not rise tomorrow. Hence, if it is irrational to believe such things, as it surely is, it cannot be because it would be contrary to our empiricist values to do so. Therefore, the irrationality of such beliefs must be explained in other ways, or the appearance of irrationality dispelled. I assume that the latter cannot be done. Therefore, to explain the irrationality of such beliefs on the kind of model we are looking for, we must suppose that there are important epistemic values which underlie our judgements of truth and falsity besides the empiricist ones.

I postulate that we are innately disposed to value those kinds of beliefs and systems of beliefs which are biologically useful to us. Perceptually and introspectively acquired beliefs clearly fall into this category. No doubt having a consistent belief system does too. But our survival also depends on our ability to anticipate nature; and this depends largely on our general beliefs and theories. It is to be expected, therefore, that we should attach a good deal of value to our store of such beliefs and theories. We should be reluctant to abandon them, especially those that are well entrenched or have served us well in the past; and if forced to do so, it is to be expected that we should try to salvage what we can from them.[8] It is also to be expected that we should strive to increase our general knowledge and theoretical understanding, and hence that we should have some epistemic values adapted to these ends.

In the following sections of this paper, I shall describe a number of processes which are evidently involved in the construction, establishment, integration and maintenance of general knowledge—in particular, the processes of generalising, normalising, testing and reconceptualising. I shall not discuss the processes involved in theory construction, or the criteria by which theories are evaluated. My aim is just to describe the basic strategies, and to identify the epistemic values they serve. I shall argue that there are four basic epistemic values which are served by these processes: viz. regularity, epistemic conservatism, corroboration and connectivity. The first provides motivation for the strategies of inductive generalisation and normalisation, and for assuming a kind of 'uniformity of nature' principle. In the area we are concerned with, epistemic conservatism exercises control over what inductive generalisations we may make. It also helps to motivate the normalisation strategy. Corroboration is the value most involved in testing general knowledge, and is therefore fundamental to the process of establishing it and determining its scope. Connectivity is a value concerned with the establishment of conceptual connections, and hence with the integration of knowledge. It motivates conceptual changes, and thus serves as a counter to our natural, and proper, epistemic conservatism.

4 INDUCTION AND EPISTEMIC CONSERVATISM

Induction is the primitive strategy we use to arrive at generalisations. Roughly, if something holds in a number of cases without exception, we are naturally disposed to think it will hold generally. What holds in most cases we know about, we are inclined to believe will continue to hold in most cases. What has rarely happened in certain circumstances, we naturally suppose will continue to happen only rarely in such circumstances. I do not think we can be much more precise in our description of the primitive strategy than this, although we are at liberty to theorise that an ideally rational being would generalise according to some more precisely defined inductive rule, such as the straight rule.

The strategy is primitive in the sense that it is unlearned. We could not have discovered it for ourselves, for we should have to have used it to do so. We were not taught it orally, because we needed it to learn at least the vocabulary of the language. For without it we should not know what any term meant, or if we did, that it would continue to mean the same. Nor is there any reason to think that we were ever conditioned to think this way. A crude disposition to generalise like this seems, therefore, to be a natural tendency. Yet, there are some generalisations we might have arrived at in this way which we should not accept, and others we find quite acceptable. So, however the process of inductive generalisation may be articulated, an important question remains concerning the evaluation of inductive generalisations.

---

[8] Recent empirical work on theory maintenance supports these general conclusions. See Nisbett and Ross [1980] Ch. 8 for an excellent review of the psychological literature on this topic.

The general principle must be that we tend to value, and therefore seek, generalisations of the kinds that proved useful to us as language and thought evolved. These generalisations will naturally be concerned with the sorts of things we think there are, and therefore, historically, with the sorts of things which were included in our primitive ontologies. Now, we seem naturally disposed to think of our environments as consisting of objects in space and time, possessing various properties, doing various things, and having various effects. Consequently, we are likely to value generalisations concerning any of the sorts of things we think belong to any of these basic ontological categories.

I assume that the ontology we intuitively work with is reflected in the language we use to describe the world. Therefore, since we have names for certain *kinds* of objects and their properties, and verbs describing certain *kinds* of actions and events, our primitive ontologies must include kinds of things, actions and events, as well as individuals of these kinds. Therefore, given the general principle that our epistemic preference should be for the sorts of generalisations that have been useful to us in the past, and the presumed utility of our ontologies, we should prefer beliefs about recognised kinds of things to generalisations which are not about recognised kinds, i.e. we should prefer generalisations which derive from conceptually and ontologically conservative descriptions of things. There is nothing sacrosanct about our intuitive conception of reality, as physics has taught us, but it is the natural base from which we must begin; and any change to it needs to be established, and its worth demonstrated, before it can be used as a basis for inductive inferences.

Normally, the logical complements of things, or kinds of things, do not belong to the same ontological categories as the things or kinds they complement—if, indeed, they exist at all. For 'Crows exist' and 'Non-crows exist' are very different kinds of claims. The first is just the claim that things of a certain kind, viz. crows, exist. The second is a much more complicated claim, because, in the sense in which it is true, it implies that there are kinds of things other than crows, and that there are things of at least some of these kinds, i.e. it is doubly existential. Interpreted straightforwardly, in the same way as the claim that crows exist, the claim is simply false. For there is no kind of thing which consists of all kinds of things which are not crows—at least not in my ontology. Likewise, the claim that non-Brian Ellis exists is, if true, just a very odd way of saying that there are individuals other than me. But if it is interpreted as the claim that there is *an individual* which consists of everything which is not me, then the claim is false. For there is no such individual.

I do not know how to distinguish between complementary terms which name genuine individuals or kinds from those which do not. I think it is a question of which we think name entities that are required for causal explanations. Certainly, causal involvement is one of our main criteria for what is physically real. But, however the distinction may be made, we all certainly make it, and it is essential to sound inductive reasoning that we should be able to. For inductive arguments must proceed from epistemically preferred descriptions if they are to provide good reasons for their conclusions. And the preferred descriptions are those which, given our current understanding of the world, are conceptually and ontologically most conservative. The irrationality of arguing inductively from the existence of white swans or yellow bananas to the conclusion that all non-black things are non-ravens derives, not from the form of the inference, which is sound, but from the irrationality of the conception of such things as non-black non-ravens.[9] It is irrational because it violates the requirement of epistemic conservatism.[10]

The epistemic values of conceptual and ontological conservatism likewise rule out the Goodman inferences. Their irrationality stems from the irrationality of the conceptions from which they proceed. To conceive of emeralds as grue is already to admit into one's ontology a property of grueness, which we could not accept without doing a great deal of violence to our conceptual framework—thus abandoning much of what we think we know for a concept of no proven worth. Our intuition, which, I suggest, derives from the value we attach to established conceptions, is to reject this description along with the inference. Obviously, we cannot reject the description of emeralds as grue on empiricist grounds alone. For the description is compatible with everything that is empirically certified. It does not, however, employ established concepts. Therefore, before we could use such descriptions as bases for inductive inferences, we should need some argument to show that such descriptions were preferable to the established ones. In the absence of such an argument, the requirements of epistemic conservatism rule them out as acceptable bases for inductive inferences.

Epistemic conservatism is not to be equated with dogmatism, or with blind adherence to traditional ways of looking at things. It is neither of these things to refuse to accept the Goodman inferences, or the existence of a white horse as evidence for the generalisation that all ravens are black. Nor is it incompatible with adventurous theorising. If someone can produce a theory with a different ontology from the usual, which is demonstrably superior, by the criteria relevant to theory evaluation, then this theory should be accepted. What it does is establish the onus of proof. Those concepts, laws and theories which have been useful to us are to be retained in the absence of compelling reasons why not.

---

[9] I do not deny the relevance of negative-negative cases to the analysis of correlations. Indeed, our natural disposition to ignore such cases is not always rational. (See Nisbett and Ross [1980], Ch. 5). What is irrational is to treat such cases as reflecting correlations between kinds, and so base an inductive argument on them.

[10] I should, in fact, be willing to accept a more general thesis of epistemic conservatism than this ontological one—something very close to Sklar's 'methodological' conservatism. For he argues, as I do, that 'Stability of belief is itself a desirable state of affairs and an end to be sought'. (Sklar [1975], p. 389.)

5 CORROBORATION

The values of epistemic conservatism restrict the range of inductive generalisations which are prima facie acceptable. The value of corroboration contributes to their entrenchment as beliefs and to determining their scope. I think Sir Karl Popper was right to recognise its importance as a value, and that he correctly perceived its role in the development of scientific theory. (See Popper [1959]).

However, Popper was wrong to distinguish, as he did, between corroboration and confirmation, because he left himself without any adequate explanation of the rational preferability of well corroborated beliefs. He cannot say that it is rational to be guided by them because they are most likely to be true. Nor can he say that having been well corroborated they are likely to continue to be. Indeed, on his own theory, we have no reason to believe that falsified theories will not always be corroborated in future, and hence no more reason to trust corroborated theories than uncorroborated ones. In my view, Popper needs corroboration as a basic epistemic value, and a theory of rational belief which recognises this value. The more often, and the more widely varied the circumstances, in which a law, theory or generalisation has been corroborated, the more rational it must be to believe it, and hence the greater its epistemic probability.

The value of corroboration, like that of epistemic conservatism, derives from the proven utility of corroborated beliefs. It is not that we know in advance that corroborated beliefs will continue to be corroborated. What we know is that they have, or would have, been successful predictors. Now the general principle from which our epistemic values may be derived is that the *sorts* of beliefs which have been useful to us in the course of evolution are the ones we are likely to value epistemically, and therefore to consider belief-worthy. So corroborated beliefs should, by this principle, be considered worthy of belief—the more so, the more often, and the more varied the circumstances, in which they have been corroborated.

The point of varying the circumstances of corroboration is to test a hypothesis more widely. If the hypothesis is that all As are Bs, and if all instances of this generalisation that have been observed are As that are Cs, and if it is plausible, on background information, that an A's being a C is relevant to its being a B, then the corroboration we have obtained may only be for the more restricted hypothesis that all As that are Cs are Bs. So, to remove this doubt about the scope of the corroboration, we must examine some As that are not Cs to see whether they too are Bs. Background information is thus obviously relevant to corroboration. For we need to know what sorts of factors might limit the scope of a generalisation to know what tests to perform. Blindly repeating tests under more or less similar circumstances is not a good way to corroborate a hypothesis. For it leaves open the possibility that the generalisation holds only in these circumstances.

On the other hand, if we thought that observation itself, or time or place of observation, might be a factor limiting the scope of any generalisation, then strong corroboration of unrestricted generalisations would be impossible. For while observing As that are Bs may make it reasonable to believe that all As that have been or will be observed are Bs, or that all As in our spatio-temporal region are Bs, we should not be able to corroborate strongly the more general hypothesis that *all* As are Bs. I assume, however, that strong corroboration of hypotheses of unrestricted scope or generality is possible, for otherwise the laws and theories of science could not be well corroborated. Therefore, we must be strongly disposed to believe, independently of corroboration, that what holds for us in our own spatio-temporal region holds generally. That is, we must operate with a sort of 'uniformity of nature' principle. Why this should be so will be considered presently.

It is impossible to say, without reference to our background theories, how strongly a hypothesis is supported by corroborative evidence. It depends how unequivocal the evidence is, how plausible the hypothesis is given our background understanding, how probable we should take the evidence to be independently of the hypothesis, how broad the scope of the hypothesis being tested, and whether there are reasons for thinking that some scope restrictions on the hypothesis being tested may be necessary.

If the degree of corroboration provided by the evidence e for the hypothesis h is measured by $P(h/e)-P(h)$, normalised to a zero to one scale,[11] then

$$1.\quad S(h/e)=P(h/e)\times P(\sim e/\sim h)$$

That is, the more probable the hypothesis, given the evidence, and the more improbable the evidence, given the denial of the hypothesis, the greater the degree of corroboration. Thus, the strongest corroborations of our hypotheses are to be obtained from their least probable consequences.

Also, if h entails g, and g entails e, then the following *chain rule* holds:[12]

$$2.\quad S(h/e)=S(h/g)\times S(g/e)$$

---

[11] I have argued in detail for this as a measure of degree of positive support in Ellis [1970]. The normalising factor required is $1/(1-P(h))$.

[12] This striking result is a convincing demonstration of the importance of test variety; and reflects favourably on the normalised scale of evidential support, since the chain rule does not hold for other measures of support.

Consequently, if the evidence we have all corroborates g, and g has been sufficiently corroborated, there is no point in seeking further evidence of this kind. To increase support for h, we must seek evidence of a different kind, viz. evidence for some consequence of h which is not a consequence of g. If h is the hypothesis that all swans are white, it is not to the point to go on examining European swans, once the hypothesis g that all European swans are white has been sufficiently corroborated. For the maximum degree of corroboration for h obtainable this way is $S(h/g)$, which, given our background knowledge of bird colour variation with geographical location, may be quite low.

These two results define and explain the importance of test severity and test variety in the process of corroboration. If a hypothesis survives sufficiently many and various tests of enough severity, it will be well-corroborated. And a well-corroborated hypothesis, I hold, is one which, other things being equal, we ought rationally to believe. Hence, there may come a point, in the process of corroboration, when the evidence for a hypothesis becomes compelling, and if it is contrary to anything else we believe, it may be necessary to revise this belief. The claims of epistemic conservatism may thus be overcome by the corroboration of hypotheses.

6 UNIFORMITY OF NATURE

It is often held that induction depends on the assumption that nature is uniform, i.e. that all things are governed by strictly universal laws. However, no principle of the uniformity of nature can be formulated which will justify our inductive practices. For it cannot be used to validate any inductive inferences, unless we can identify those local regularities which instantiate universal laws. Moreover, the assumption itself needs justification, since it is not a priori, nor even generally accepted. The Ancient Greeks, for example, expected to find perfect regularity only in the heavens, and in the abstract world of the Forms. And many people today would exempt human actions from the ambit of universal laws. So the assumption is not justified either way. But even if it were, we do not appear to need it. The Ancient Greeks argued inductively, though they did not make this assumption.

Nevertheless, there is something to the suggestion that belief in the uniformity of nature supports our inductive practices. For we seem naturally to be disposed to search for regularities or uniformities in nature, and to see any deviations from the regularities we think we have discovered as being superimposed by extraneous forces. It is not that nature, as we find it, is uniform, even within the range of our observations. But we are inclined to assume it is, nevertheless, and to conceive of things as instantiations of general regularities, distorted though they may be by their peculiar circumstances or histories. We seem thus to be natural *regularity seekers*. We are not content with local generalisations. Whatever holds locally, we are naturally inclined to suppose holds generally, unless there are special reasons (such as those the Greeks had, or libertarians have) to think otherwise. Newton summed up the attitude in the third of his 'Rules of Reasoning in Philosophy':

> The qualities of bodies which admit neither intensification nor remission of degrees, and which are found to belong to all bodies within the reach of our experiments, are to be esteemed the universal qualities of all bodies whatsoever. (Thayer ed. [1953], p. 3)

To explain this attitude, which presupposes the uniformity of nature, I assume that the drive to discover regularity is primitive. It is not that our inductive reasoning is justified by our belief in the uniformity of nature. It is rather that the epistemic values which determine the patterns of our inductive reasoning naturally incline us toward this belief. Belief in the uniformity of nature cannot justify our inductive inferences; nor can that belief be justified inductively. But there is an underlying attitude, which does much to explain both, viz. the primitive desire to discover regularity in nature.

The value of regularity manifests itself in a number of ways. Most obviously, it inclines us to believe that generalisations which are found to hold locally also hold generally. But more importantly, it drives us to search for deep or *underlying* regularities in phenomena, where no such regularities are apparent, or if apparent, then not unblemished.

The search for regularity has led to the development of a number of strategies. Of particular importance are the strategies of *normalisation* and *idealisation*. The first of these is a primitive strategy for concept formation; the second, a basic strategy for theory construction. The normalising strategy is primitive, at least in the sense that it is pre-verbal. For without it we could not acquire knowledge of the concepts or grammar of a language. The strategy of idealisation is probably not primitive in this sense. It appears to be an extension of the strategy of normalisation—one which is useful when the conditions for the applicability of the normalising strategy are not satisfied.

To normalise is to distinguish between normal and abnormal cases in order to formulate or defend some generalisation, and to isolate any real or apparent counterexamples. A normalised generalisation is then seen as being the statement of some *underlying* regularity; and the counterexamples are seen as being as more or less superficial distortions produced by special conditions or circumstances.

Philosophers have not paid much attention to the strategy of normalisation which is standardly used in connection with empirical generalisations. Yet the normal/abnormal distinction, and its many variants, is as fundamental and pervasive as any in our language. We apply it to people, stars, modes of behaviour, speech patterns, chemical reactions, circumstances, samples, and just about everything else. Indeed, to conceive of things in terms of norms and exceptions is normal practice, and it seems to come quite naturally to us. No theory of human rationality can be satisfactory, therefore, if it gives no account of it.

Evidently, the disposition to normalise is primitive. For without a readiness to distinguish between normal and abnormal cases, and thus maintain generalisations in spite of exceptions, we could never learn to speak. Most of our empirical concepts depend for their identity on the set of normalised generalisations we accept concerning them. Thus our concept of man is not defined by the set of strictly universal generalisations we accept about men. There are too few such generalisations, and they are not specific enough to distinguish men from many other animals. What distinguishes men at the perceptual level, which is the level at which our concept is acquired, is what they *normally* look like, what capacities they *normally* have, what features are *normally* present, and so on. And the identity of our initial concept of a man must depend on these. The same is true of most of the other empirical concepts we need to acquire a knowledge of language. Therefore, the normalising strategy must be primitive enough for us to be able to operate with it at this level.

7 CONNECTIVITY

The desire to understand things is the main motivation for scientific enquiry, and we should expect there to be several primary epistemic values concerned with this activity. Here I shall focus on one which I think is important. I call it the value of connectivity, because it is concerned with the establishment of conceptual connections, and hence with the theoretical integration of knowledge.

The inductive practices I have considered so far have been more or less straightforward inferences from samples to populations. Those involved in the establishment of conceptual connections are much less straightforward. Indeed, they include the full range of strategies described by Imre Lakatos in his 'Proofs and Refutations' (Lakatos [1963]).

The strategies of normalisation and idealisation enable us to identify certain underlying regularities in nature, and to define the irregularities which remain to be accounted for. The acceptability of the norms and ideals we thus arrive at depends on how well we can explain what is not normal or not ideal relative to these concepts and principles. Hence our normalising and idealising strategies cannot in themselves satisfy our desire to discover uniformity or regularity in nature; they merely set an agenda for doing so.

To explain why some As are Bs, given that As are not normally or ideally Bs, we must try to discover some characteristics of the As, or the circumstances in which we find them, which are sufficient to account for their B-ness. So, minimally, we must look for generalisations of the form, 'All As that are Cs are Bs'. However, not every acceptable generalisation of this form is explanatory.

For example, 'All As that are Bs are Bs' is not. To be explanatory, the condition C cannot be more than *contingently* sufficient for the B-ness of As.

There may, however, be several contingently sufficient conditions for the B-ness of As. Thus, all As that are Ds may be Bs, as well as all As that are Cs. Moreover, there may be As that are Bs which are neither Cs nor Ds. So we may have to say that some As are Bs for other reasons. Furthermore, the explanatory hypotheses we arrive at may themselves turn out to have exceptions, and so have to be normalised. For example, we may find that C is only normally a sufficient condition for an A being a B. And then, no doubt we should want to know what sometimes prevents C from having its normal effect. The search for explanations of the exceptions to our normality generalisations, or to the framework principles postulated as governing the behaviour of the idealised theoretical entities of science, may thus lead to many complications.

Proliferating subsidiary explanations in this kind of way is intrinsically unsatisfying, and if it seems that there are too many reasons why an A may be a B, then we are likely to think that we have not got to the heart of the matter. For example, we may think that the appearance of diversity amongst the reasons for B-ness is superficial, or that the allegedly similar effects are fundamentally dissimilar. In extreme cases, we may think that the problem is generated by the norms or ideals we have adopted, and so seek some other way of conceptualising the subject matter of our enquiry.

Our search thus appears to be for conditions which are not only sufficient for the effects we wish to explain, but also, if possible, necessary. And persistent failure to discover conditions which are both necessary and sufficient for a given effect can be a powerful incentive to reevaluate the conceptual framework within which the problem arises.

The discovery of such conditions for things is evidently intellectually satisfying. So, presumably, some epistemic value attaches to the knowledge of them. This is the value I call 'connectivity'. The source of this value, I suppose, is that such knowledge forges the conceptual links we need to build up an adequate conceptual framework for interpreting reality, i.e. it is an important contribution to our *understanding* of things. If we know that A is a necessary and sufficient condition for B, then the knowledge that something is a B is not isolated, but linked to the knowledge that it is an A. Hence, if we can discover what makes something an A, we may also find out what makes it a B. Or perhaps we shall find that there is some common cause for both A-ness and B-ness. Either way, the knowledge that something is a B may be located in a systematic framework, and so better understood.

The biological advantage of establishing such connections between items of knowledge is the increased efficiency and precision of the knowledge system it generates. If we know that As are normally Bs, then that may sometimes be good enough. But if we know that all and only As that are Cs are Bs, then the knowledge is more precise, and we have at least the beginnings of a theory of B-ness. Moreover, if we know this, then the fields of evidence for C-ness and B-ness are both increased. For now, any evidence that an A is a B is immediately evidence that it is a C, and conversely. Consequently, the establishing of such a conceptual tie between C-ness and B-ness expands the range of inferences we can draw from the evidence we have, and consequently increases the efficiency of our knowledge system.

I postulate that the desire to know what makes something the case, in the sense of knowing necessary and sufficient conditions for it, is primitive. The interest we have in knowing this is not just the interest of controlling or anticipating nature, although this may help to explain why we have it. For the desire to know the necessary and sufficient conditions for things is manifest even where there is no possibility of control, or the events we are seeking to explain are unlikely to recur. The interest is rather in satisfying our intellectual curiosity, our desire to understand what makes things so.

This desire is partly satisfied by the discovery of conditions which are either necessary or sufficient, for such discoveries also increase the range of inferences we are able to draw. But conditions which are both necessary and sufficient appear to be especially significant to us, and we are often willing to sacrifice content, and to modify or refine our concepts to establish such connections, as Lakatos' history of Euler's theorem amply demonstrates. (*ibid.*) If all As are Bs, but some Bs are not As, then A-ness does not adequately account for B-ness. On the other hand, A-ness cannot *by itself* account for B-ness if there are any As that are not Bs. Therefore, we can have an adequate theory of B-ness if we can find some condition which is both necessary and sufficient for B-ness.

The desire to discover necessary and sufficient conditions for things is thus a very important ingredient of our desire to understand things. So it is not surprising that we should sometimes be willing to modify our concepts, or to reduce the empirical content of our assertions, in order to create such connections. Consider the Lakatosian history of Euler's theorem. (*ibid.*) As the classroom discussion opens, the concepts of polyhedron, polygon, face, edge and vertex are all somewhat amorphous normality concepts. By the end, all of these concepts have been shifted or focussed, various conceptual links have been established (provisionally, at least) between them, and Euler's theorem has been embedded in a comprehensive theoretical network.

From a Popperian point of view, the changes which have occurred are content-reducing. For the tighter the conceptual links become in the process of articulating the theory, the more the propositions asserted take on the character of conceptual truths, and the less falsifiable they become. Counter-examples become not only more difficult to find, but more difficult to conceive. We are of course better able to predict the properties of polyhedra, given the theory, than we were before, because we had to learn how to apply the theory in the process of refining and developing it. The theory was improved, however, more as an instrument for conceiving and analysing polyhedra, and so predicting their properties, than as a set of independently testable assertions concerning them. And the scientist, like the carpenter who acquires a better set of tools, finds it more difficult to blame his or her equipment if things go wrong.

The value of connectivity is a basic value in theory development. For the establishment of new conceptual ties inevitably increases our understanding of things. I do not say that this is all that is involved in improving our understanding, but forging new conceptual links is at least an important part of it. The value is one which, more than any other, leads to conceptual innovations, and is therefore a counter to our natural (and proper) epistemic conservatism. It is also a value which is often opposed in its effects to the aim of making our theories more testable. The more tightly our knowledge system is integrated the less vulnerable to empirical refutation many of its elements will become.

8 SOLVING THE PROBLEM

To solve the problem of induction, it will be necessary to develop in detail the kind of epistemology I have sketched here, and show that it is adequate to explain, and, in so far as they are rational, justify our scientific inductive practices. I do not claim to have made more than a start on this. The list of epistemic values I have given is certainly incomplete, and the values contained in it are probably not all equally fundamental. For example, there is reason to think that there is a value of simplicity, related to economy of thought, which is at least as fundamental as any that I have so far mentioned. There is also, as I have argued elsewhere (Ellis [1980]), a value of objectivity, which is related to the social functions of knowledge. On the other hand, I should not be surprised if some of the values I have listed could be shown to be dependent on others. Henry Krips, for example, has argued that epistemic conservatism is not an independent virtue. (See Krips [1982].) I am convinced, however, that this is the right approach. To solve the problem of induction we need to define a biologically plausible ideal of rationality on which our inductive practices can be explained. I have assumed that, biologically, we are value-driven decision systems. If this is right, then what we must have is a values-based epistemology. The aim of this paper has been to outline such a theory.

An ideally rational being, I hold, is one that is at least: consistent; has a natural tendency to generalise widely, but is epistemically conservative in the generalisations it makes; seeks corroboration for its beliefs, and, other things being equal, believes more strongly in those that are better corroborated; is driven to seek underlying regularity in nature, and so conceive of things in terms of norms and exceptions, or, by extension, in terms of ideal and actual states of affairs, and finally; strives to increase its understanding by seeking to establish conceptual connections. An ideally rational being must also be supposed to have a belief system which is perfectly integrated according to its own standards of rationality.

But how might all this solve the problem of induction? Let us suppose that we have at last developed a values-based epistemology which is adequate to explain our scientific inductive practices, i.e. a theory on which our inductive practices either turn out to be rational according to the theory, or not rational, but for reasons we are persuaded are correct. Our inductive practices are now, by hypothesis, embedded in an adequate theory of human rationality. Hence we can now say what makes them rational. They are rational to the extent that they are the practices of an ideally rational being. But then, you may ask, why should we be rational in this sense? Why should we accept the value system of such a being? To this my answer is 'Because you are human'. For these *are* your epistemic values, and you cannot, like a god, step outside your value system to judge whether or not it is rational to have them.

La Trobe University,
Melbourne, Australia

REFERENCES

ALSTON, W. P. [1978]: 'Meta-ethics and Meta-epistemology', in A. I. Goldman and J. Kim (eds.) *Values and Morals*. Dordrecht, Reidel, pp. 275–297.

CLENDINNEN, F. J. [1982]: 'Rational Expectation and Simplicity', in R. McLaughlin (ed.) *What? Where? When? Why?*. Dordrecht, Reidel, pp. 1–25.

COHEN, L. J. [1981]: 'Can Human Irrationality be Experimentally Demonstrated?' plus 'Open Peer Commentary', *Behavioural and Brain Sciences*, 4, pp. 317–370.

COHEN, L. J. [1986]: *The Dialogue of Reason: An Analysis of Analytical Philosophy*. Oxford, Clarendon Press.

DANCY, J. [1982]: 'Intuitionism and Meta-epistemology', *Philosophical Studies*, **42**, pp. 395–408.

ELLIS, B. D. [1965]: 'A Vindication of Scientific Inductive Practices', *American Philosophical Quarterly*, **2**, pp. 296–304.

ELLIS, B. D. [1970]: 'Explanation and the Logic of Support', *Australasian Journal of Philosophy*, **48**, pp. 177–189.

ELLIS, B. D. [1979]: *Rational Belief Systems*. Oxford, Blackwell.

ELLIS, B. D. [1980]: 'Truth as a Mode of Evaluation', *Pacific Philosophical Quarterly*, **61**, pp. 85–99.

GOODMAN, N. [1954]: *Fact, Fiction and Forecast*. London, Athlone Press.

KRIPS, H. [1982]: 'Epistemological Holism', *Studies in the History and Philosophy of Science*, **13**, pp. 251–264.

LAKATOS, I. [1963]: 'Proofs and Refutations', *British Journal for the Philosophy of Science*, **14**, pp. 1–25, 120–139, 221–245, 296–342.

LEVI, I. [1967]: *Gambling with Truth*. London, Routledge and Kegan Paul.

Levi, I. [1980]: *The Enterprise of Knowledge*. Cambridge, Mass., M.I.T. Press.

Nisbett, R. and Ross, L. [1980]: *Human Inference: Strategies and Shortcomings of Social Judgement*. Englewood Cliffs, N.J., Prentice-Hall.

Popper, K. R. [1959]: *The Logic of Scientific Discovery*. London, Hutchinson.

Pugh, G. [1978]: *The Biological Origin of Human Values*. London, Routledge and Kegan Paul.

Ross, W. D. [1939]: *Foundations of Ethics*. Oxford, Clarendon Press.

Salmon, W. C. [1982]: 'Further Reflections', in R. McLaughlin (ed.) *What? Where? When? Why?* Dordrecht, Reidel, pp. 231–280.

Sklar, L. [1975]: 'Methodological Conservatism', *Philosophical Review*, 84, pp. 374–400.

Thayer, H. S. ed. [1953]: *Newton's Philosophy of Nature*. New York, Hafner.

Brit. J. Phil. Sci. 39 (1988), 161–182 *Printed in Great Britain*

A Theory of Probability

T. V. REEVES

ABSTRACT

This paper argues that probability is not an objective phenomenon that can be identified with either the configurational properties of sequences, or the dynamic properties of sources that generate sequences. Instead, it is proposed that probability is a function of subjective as well as objective conditions. This is explained by formulating a notion of probability that is a modification of Laplace's classical enunciation. This definition is then used to explain why probability is strongly associated with disordered sequences, and is also used to throw light on a number of problems in probability theory.

1 *Sequences and Probability*
2 *Generating Sources and Probability*
3 *A Theory of Probability*
4 *Applications*

I SEQUENCES AND PROBABILITY

The extensive use of probability calculus in the making of accurate statistical predictions is a persuasive fact in favour of the argument that the notion of probability has its basis in objective conditions. This fact alone, perhaps, is enough to explain the considerable backing objective theories of probability receive, a backing that appears in a variety of forms. One of the most elementary constructions about which statistical predictions can be made is the sequence. Sequences are of fundamental concern because they are readily measurable manifestations of single 'sources of probability' that have been set to work to generate outcomes. A sequence can be conceived of as a configuration of things such as numbers, objects, or events etc. Sequence frequency parameters can be readily defined as various enumerations made upon sequence features, and a given sequence can be counted, measured and recorded with an objectivity that is about as contentious as a stock take. Although we can define a variety of mathematical constructions based on sequence parameters, a kind of reductionism applies here in as much as these constructions derive from a sequence of objects or events that can effectively be reduced to a representation which is fundamentally a digital pattern. In this sense a discussion of sequences can be developed using a notation that does not employ probability. If we do introduce and relate the concept of probability to this relatively prosaic world, it tends to give rise to unnecessary intellectual discomforts; The clear cut, black and white world consisting of the 'yes' and 'no's' of digital representation are much easier to wrap the mind round than the shades of grey introduced by the varying degrees of 'perhapses' and 'may be's' of uncertain meaning found in probability theory. However, these discomforts can be avoided by restricting a discussion of sequences to the issue of form and configuration only. If one does this, and also neglects the question of the meaning of form and configuration one find that sequences yield readily to mathematical analysis. This is something that RF theorists and configurationalists have demonstrated by carrying out valid mathematical discussions of sequences before there is any mention of probability. It is true that such discussions sometimes wind up by defining probability, almost as an after thought, as some construction (usually an RF) based on a certain sequence type, but although this step may be philosophically contentious, it is not difficult to understand the temptation of this sort of configurational treatment of probability; If the identification of probability with some configurationally defined parameter is correct, then it looses much of its mystique by becoming the object of a mathematical rather than philosophical investigation; probability, we are effectively told, reduces down to the mundane issue of sequence form. But, whatever we think of the philosophy of RF/configurationalist accounts of probability, the mathematics contained therein, although amounting to little more than theorems on sequences, is valuable—valuable because it deals with an area that probability calculus appears to make predictions about: namely the sequence. It is the apparent ability of probability calculus to make predictions about such 'hard digital entities' as sequences which convinces many that a given probability must be rooted in objective conditions. Perhaps, then, an alternative theory of sequences that does not employ probability as a predicate, and of a clear cut character, that shares a common ground with probability in its ability to make predictions about sequences, may throw light on the nature of probability itself.

But there is a problem here that becomes apparent when we consider the nature of the predictions about sequence form that probability calculus makes. Consider, for example, a sequence generated by a source that has a fixed probability of returning a particular outcome (which means that the probability of the outcome does not change during sequence generation). Now, let us define a 'local configurational type' as a binary pattern consisting of a pattern formed by the relative positioning of outcomes over a limited number of contiguous trials of the generating source. If we generate a long sequence then a particular local configurational type (which must consist of a lot less trials than the total sequence) will appear at a variety of positions in the sequence. If we arbitrarily select a series of positions in the sequence (called a reference class of positions) we may find that we have selected some of the positions associated with the local configurational type in question. Now, probability calculus predicts that for long sequences there is a high probability of the relative frequency of a configurational feature over a reference class being approximately equal to the calculated probability of the feature. In general terms this fact can be expressed as:

$$\text{PROB [(RF of feature over reference class)} \approx \text{PROB(feature)]} \rightarrow \text{one,}$$
$$\text{as sequence length increases.}$$

i.e. As the length of a sequence increases, the probability of RF values equalling their corresponding probabilities tends to one.

Because of the high probability associated with conditions where RF values are approximately equal to their corresponding probabilities, we think of them as most representational of probabilistic generating sources. These conditions are assumed when practical statistics is applied to estimate frequencies, and therefore one could refer to sequences that have these conditions as 'statistically valid'. For infinite sequences statistical validity is a certainty, and for this reason there may therefore seem some justification in asserting a relation something like:

$$\text{PROB(feature)} \equiv \text{Limiting RF of feature.}$$

This is the sort of relation between probability and relative frequency that an RF approach to probability affirms. But the standard problem with such an approach is that though probability calculus may predict with a high probability that for a long sequence the RF of a feature over any reference class is equal to its probability, it is simply not logically possible for a real sequence to be statistically valid on all reference classes. A given sequence, by logical requirement, contains certain ways of selecting reference classes whose RFs diverge considerably from the 'predicted' values. This means that it is not logically possible to simulate in one sequence all the high probability predictions of probability calculus concerning statistical validity with a purely configurational/RF theory of sequences. Thus, any attempt to literally interpret probability as the Relative Frequencies predicted by some kind of sequence calculus, is obliged to pay the price of accepting that such a calculus cannot consistently predict a real sequence that embodies at once, every RF characteristic that probability calculus predicts as having a high probability. This major philosophical problem of RF theories of probability is usually tackled by limiting the allowed reference classes or 'gambling selection schemes' in such a way that consistency is saved, and yet to prevent this limit being so severe that the calculus fails to produce useful results (see, for example the work of Church and Wald). But, procedures of this kind effectively write off certain anomalous predictions of probability calculus as invalid; It does not accept the fact that probability calculus predicts that all reference classes, regardless type, have for one sequence, a high probability of displaying statistical validity. It would be nice if we could somehow accept and yet explain this strange feature rather than disallowing it. Probability calculus does not make clear cut, unambiguous predictions of a 'yes' or 'no' quality—all such predictions are prefixed with a probability. This ambiguous nature saves the calculus from contradiction; It gets away with what would otherwise impose contradictory demands on an ordinary sequence calculus because a high probability is not a logical bind. But if we do not identify probabilistic predictions with the binding assertions of sequence calculi, the price we have to pay for this is that these predictions remain enigmatic and of uncertain meaning.

Another characteristic of probabilistic prediction that any purely configurational theory of sequences would find difficulty in simulating arises from the following consideration. Imagine we have a sequence of 1s and 0s generated by a source with a fixed probability of P_1 of producing a 1 and a probability of P_0 $(=1-P_1)$ of producing a 0. Let us arbitrarily select a particular sequence and ask ourselves what is the probability of it being generated. For example, what is the probability of the following sequence consisting of ten trials?

1011010111

By requiring that the source has a fixed probability of generating a 1, we are in effect requiring that the probability of this outcome is independent of its surrounding sequential context. Thus, the probability of the above configuration is:

$$P_1P_0P_1P_1P_0P_1P_0P_1P_1P_1=P_0^3\,P_1^7$$

It is quite clear from this example that if we rearrange the 1s and 0s but keep their respective frequencies constant, the probability of the new configuration has the same value as that of the above configuration, namely $P_0^3P_1^7$. By obvious extrapolation it is apparent that any two sequences of the same length and possessing the same number of 1s and 0s have the same probability of being generated. This of course means that any intuitively 'irregular' sequence has the same probability of being generated as regular repeating sequences such as 10101010. . ., provided the two sequences have the same frequencies of 1s and 0s. If the latter condition is maintained and we increase the lengths of the irregular and regular sequences but keep them of the same length, clearly the ratio of the probabilities of the two sequences remains as one, even though we may extend the sequences to infinity. Thus, as far as their probabilities are concerned, probability calculus does not make a distinction between 'regular' and 'irregular' sequences, whether we are talking about finite or infinite sequences. Consequently, we have the strange result that a large or infinite 'irregular' sequence generated by a probabilistic source, has a probability of existence no greater than the minute probability associated with a regular sequence with the same frequencies of 1s and 0s. It is difficult to see how this result could be paralleled in a sequence calculus intended to simulate probabilities with relative frequencies; One would expect such a calculus to be constructed with the very purpose of excluding absolutely the 'regular' sequences we consider to be so untypical of probability. For example, this is, in fact, precisely what Von Mises does with his axiom of randomness. But once again we see here how a configurationalist interpretation of probability requires logically binding assertions where a natural approach to probability calculus arrives at more tolerant conclusions.

. . . . .

Sequence calculi, which are essentially theories of digital configurations, can not give complete account of the grey areas of probability. RF/configurational interpretations of probability inevitably impose logically binding 'yes' and 'no's' where probability calculus is more ambiguous. However, it is true that sequence calculi can usefully simulate many of the results of probability calculus. In an unpublished work I have attempted to address this area by defining a configurational concept of disorder and applying it to sequences. Using the related 'disorder theory' one can derive, as others have done in different ways, such conclusions as Bernoulli's formula. One important understanding of this theory is that 'disordered' configurations tend to have a high 'statistical weight', i.e. the percentage of the total possible configurations that are 'disordered' is a majority percentage. This understanding throws light on why it is possible to simulate probability calculus with a sequence calculus; the high statistical weight that one would expect to be associated with 'disordered' sequences will mean that this class of sequences has a high probability. Thus, any sequence calculus that effectively identifies the class of disordered sequences (or 'random' sequences, as they are usually called) will be able to predict the features of a very probable outcome of probabilistic sources. This correlation of disorder and high probability may prompt the identification of disordered sequences with probability, but to make this identification without paradox requires the absolute elimination of some of the anomalous predictions of probability calculus.

2 GENERATING SOURCES AND PROBABILITY

A 'static distribution' is a distribution or sequence that has no obvious connection with a probabilistic generating source, but just exists, like for example a random star distribution. In the latter case it is very natural to talk about the probability of a star being at a location, but on closer consideration we find one or two anomalies in dong so. How can we talk about the probability of a star occupying a location when it seems that the objective reality of the distribution requires unambiguously that a position is either occupied by a star (cf $P=1$) or it isn't (cf $P=0$)? Furthermore, the state of the distribution as a whole consists of an unambiguous and non-contradictory configuration of stars, and therefore, in this particular objective connection it is difficult to make sense of those aspects of probability calculus that give contradictory outcomes and configurations equal weight (see previous section). If it is right to talk of a probability in this situation and if this probability resides in objective conditions, where are these conditions and what is their nature? Is the probability considered to pertain to the source that originally generated the distribution, or to the ongoing dynamics that moves the stars about? (But let us note here that there seems to be no logical reason why either generating source or a dynamics should be concerned here; perhaps the distribution has no dynamics and is infinitely old, and therefore had no generating source either.) Similar points can be made in connection with probabilistically generated events that have become history; is it right to refer to the probability of a 'happened' event? As for static distributions, the unchangeable and unambiguous reality of historical sequences of events, makes little sense of the ambiguities introduced by probability calculus. For static distributions and historical sequences the configurations have somehow become 'fixed', and therefore it may seem unacceptable to see a probability as somehow residing in the distribution or sequence themselves. Both cases do not appear to possess in any shape or form the 'indeterminate' middle ground suggested by probabilities—they are situations that are black or white, true or false. The immobility of these situations contrasts with the strange dynamic processes at the heart of probabilistic sources that generate sequences and distributions; what a source will eventually do does seem to be in some kind of objective melting pot until it has done it. i.e. The lack of predetermination toward particular outcomes seems objectively present in the source itself. Perhaps, then, sources that generate sequences and distributions are a more promising receptacle for objectively embodying the strange balances of opposite outcomes suggested by the notion of probability (although this leaves us with a problem for the star distribution where there is no logical mandate requiring a generating source).

But, identifying probabilities with the mysterious influences at the heart of dynamic processes gives rise to its own problems. To illustrate this, imagine that we have two coins, one of which is unbiased and the other biased toward heads. Imagine that we have a third coin which when thrown is used to decide which of the other two coins to throw—For example, if the third coin landed tails up the unbiased coin was thrown, otherwise the biased coin is thrown. Thus, at the throw of the third coin (i.e. the 'control' coin), either the biased or unbiased coin would be thrown, depending on the outcome of this throw. Now, given this situation, it is a meaningful question to ask what is the probability of getting a heads if either of the biased or unbiased coins are thrown. To answer this question let P1 $(=\frac{1}{2})$ be the probability of getting a heads when the unbiased coin is thrown, and let P2 $(>\frac{1}{2})$ be the probability of getting a heads with the biased coin. If the control coin is unbiased then the probability of its throw being followed by a toss of the unbiased coin which results in a heads is given by $P1\times\frac{1}{2}$. Similarly, the probability of getting a throw of the biased coin resulting in the outcome of a heads is $P2\times\frac{1}{2}$. Consequently, the probability of getting heads on the throw of either the biased or unbiased coin is equal to $\frac{1}{2}(P2+P1)$. But there is a paradox here; How can it be that on the throw of, say, the unbiased coin, the probability of getting heads is $\frac{1}{2}(P2+P1)$ when it is clearly P1? It seems that the probability of getting a heads must be either P1 or P2, so where does $\frac{1}{2}(P2+P1)$ come from?

Is there, for this three coin system, such a thing as a single, unambiguous, and absolute probability of getting heads? Perhaps, because the throwing of the biased and unbiased coins is conditioned on the result of the throw of the control coin, the value $\frac{1}{2}(P1+P2)$ appears to be more absolute than P1 or P2. But, in the description of the three coin system, it is not necessarily ruled out that the throwing of the control coin itself is governed by a higher level coin that determines whether the control coin conditioning the throws of the unbiased and biased coins is one of two possible coins of different probability characteristics. i.e. A 'high level' control coin is thrown, the result of which selects one of two lower level control coins to be thrown, which in turn select the biased or unbiased coins at the bottom of the 'hierarchy'. Thus, because the throw of the control coin itself is conditioned on a probability, $\frac{1}{2}(P1+P2)$ is not necessarily the 'absolute' probability of heads. In fact it is clear that it is possible to develop the control coin hierarchy to the n^{th} degree: At each level in this hierarchy there are two control coins, where one of the pair is selected to be thrown depending on the outcome of the throw of a control coin one level higher up in the hierarchy. This latter coin is in turn itself selected from one of two coins, depending on yet another layer in the hierarchy, and so on to n levels. Because there seems to be no principle limiting the value of n, there is, therefore, apparently no obligation ensuring that it is actually possible to reach the top of the hierarchy. If this is the case, then any dynamic conditions to which a specific probability might pertain (like, e.g., a throw of a particular coin) are themselves the subject of a probability. Thus, all stated probabilities will be relative probabilities in as much as they can only be stated with respect to dynamic conditions, that because they are only probable, may not necessarily hold. For the three coin system it appears that $\frac{1}{2}(P1+P2)$ is the absolute probability of heads because the throw of the control coin was presented as a condition that held. But, just because it was omitted from the description of the three coin system, it does not mean to say that the throw of the control coin itself is not an outcome near the bottom of a 'control coin' hierarchy. Perhaps this could be explained by considering the relevant probabilities to only to become 'absolute' probabilities as the system unfolds itself and the conditions to which these probabilities pertain come into force and therefore are themselves no longer 'probable'. Thus, on this approach, $\frac{1}{2}(P1+P2)$ is the relevant probability of heads when the throw of the control coin is, objectively speaking, a decided outcome, whereas the probability of heads becomes P1 or P2 after the throw of the control coin, when the choice between the two other coins has become a decided matter. Seemingly, then, during the operation of the system the probability of heads has changed from some value through $\frac{1}{2}(P1+P2)$ to P1 or P2. If this is true of this system, then perhaps it is also true of other systems (like for example a simple toss of a coin), and the probability of an outcome may change with changing conditions up until the moment it has become written into history.

By identifying the notion of probability with a systems dynamic characteristics and properties there is a subtle shift of emphasis that means we do not talk so much about 'The probability of heads' but instead 'the probability of a source generating heads'. i.e. The probability does not belong so much to the outcome itself as it does to the systems and subsystems generating the outcome. If probability is to be identified with something as fickle as a systems dynamic properties such an approach does have consequences that from a gambler's point of view are intuitively unnatural; I would submit that a gambler, who, for obvious reasons, will be more interested in actual outcomes rather than mysterious dynamic properties, is going to want an answer to questions like, 'What is the probability of this outcome?'. The practical outlook forced upon the gambler means that he is likely to see probability as more bound up with outcomes rather than the complex of dynamic conditions that generate them. As an illustration consider a shuffled pack of cards: What is the probability of removing from the top of such a pack the ace of diamonds? It is natural to reply 1/52. But, is this reply correct if the pack is already shuffled? If we are looking at a static pack of cards it seems to us that the matter of which card is on the top has been decided before we make the draw. If we consider probability to be a dynamic 'tendency' toward generating outcomes, then it is difficult to see any such thing in a static pack of cards, where the enigmatic 'tendency' appears to have been replaced by an unambiguous outcome. In reality the ace of diamonds is on the top of the pack or it isn't—there is no objective middle ground that corresponds to an objective probability. This lack of middle ground in reality contrasts with the subjective point of view taken by a gambler who removes a card from the top of the pack and then looks at it; as far as the gambler is concerned the matter remains undecided until he is able to see the top card. The pack of cards could have been shuffled a hundred years ago, or it may be in the process of being shuffled, but for the gambler, whether he is thinking about what the top card is, or is going to be, his knowledge of the situation is equally undecided in both cases. For this gambler the outcome is subjectively speaking, in the melting pot right up until the last moment. There is a sense in which 1/52 really is the probability of finding an ace of diamonds on the top of the pack, as it makes little difference to the gambler's view of the matter whether the pack is static or is being shuffled. In fact we may even go as far as saying that the gambler's view of things is not changed if the pack has never been shuffled, or actually does not possess the capability of being shuffled (perhaps the cards are glued together—barring the top card, of course). Although we may feel that a static pack of cards, or a pack that can't be anything other than static (e.g. the cards are glued together), has lost some kind of inherent objective indecision or has no potential (propensity?) for possessing such, to the gambler who is making a one-off decision and knows no more than that the ace of diamonds is one out of fifty-two cards, the dynamic conditions of the pack or its dynamic propensities make little difference to his view. As far as the gambler's knowledge of the outcome is concerned, the conditions of the pack and/or its propensities do not figure, and therefore any changes in these will not be relevant to his view of the probability.

If we return to the three coin system and consider it from the point of view of a gambler-observer similar issues arise. For a particular outcome an observer may not know about any changes of condition that crop up during the trial of the system—e.g. He may not know how the control coin falls. For the observer who only knows the outcome to be a product of the three coin system, $\frac{1}{2}(P1+P2)$ is the only value he is able to assign to the probability of heads. Even though during a trial the control coin must be thrown and its result written into history, for the observer who is only able to relate the outcome to the three coin system, that outcome could just as well be the 'direct' result of a probability of $\frac{1}{2}(P1+P2)$. In fact even if the trial is complete, up until the observer is actually aware of the result, the outcome is virtually speaking the subject of a direct probability of value $\frac{1}{2}(P1+P2)$. The observer may have an intellectual appreciation of the fact that a given outcome must be either generated by the biased or unbiased coins, and he may understand that the throw of the control coin itself is conditioned on a probability. But, I submit, though he may appreciate these things, if he is an observer with a vested interest in the outcome, the value $\frac{1}{2}(P1+P2)$ will be the quantity that interests him the most—not P1 or P2 because he does not know which of these apply, and not any probability values derived using conditions further up the control hierarchy because the throw of the control coin is a given, rather than probabilistic fact.

However, we must of course realise that a particular outcome may be known by an observer to be the result of the throw of a definite coin. If this was the case, the observer would be able to map the result to different conditions—e.g. he would know the outcome to be a product of, say, the biased coin, in which case he is likely to claim that heads has a probability of P2. But what has happened to $\frac{1}{2}(P1+P2)$? This value was derived by assigning a value of $\frac{1}{2}$ to the probabilities of the biased and unbiased coins being thrown as a result of these throws being conditioned on the throw of the control coin. But as far as our observer is concerned the outcome is known to be a product of the biased coin and it therefore seems unnatural to assign a probability to the throws of the biased and unbiased cons under these subjective conditions. Thus, it is difficult to resist the feeling that $\frac{1}{2}(P1+P2)$ is the value of the probability of heads for the observer who only knows enough to map the outcome to the trial of the total system, whereas a probability of P1 or P2 applies if the observer is able to map the outcome to a single coin.

We appear to have a situation where the same outcome maps to two sets of conditions with different 'probabilities', and the resulting duality gives every appearance of having a connection with the observer's view. The problem here is in fact a special case of a more general problem that is sometimes referred to as the 'conditionality' of probability; that is, in many instances it is possible to assign two or more probabilities to an outcome in a way that depends on how a situation is described. This is demonstrated in the example of the three coin system where the probability of heads seems to depend on whether or not the observer knows which coin is directly associated with the final result; because a description either conveys knowledge or represents knowledge already held, one would expect a description to have a correlation with an observer's knowledge of a situation. Therefore, if probabilities are dependent on an observer's view, then one would expect them to be effected by the knowledge embodied in the observer's descriptions of this situation.

It does seem that probability has some sort of connection with the observer's level of knowledge, otherwise it is unlikely that a priori and subjective theories of probability would have developed in the first place; any philosophy of probability ought to include in its scope an answer to the question of why it is that, unlike other theoretical concepts such as points, lines, force, mass etc, probability has so often been linked with the subjective world. That probability is not unconnected with the subjective realm is detected when an observer claims to have 'certain' knowledge of something; if an observer knows with certainty what an outcome will be or was, then it seems unnatural that a probability should be associated with that outcome. It is an intuitive contradiction for probabilistic events to be certainties for an observer. There are other cases where probability has been linked with levels of knowledge. For example, we speak of Heisenberg's 'uncertainty' principle. Here, the concept of uncertainty, with its obviously subjective connotations of unsureness and lack of knowledge, is associated with the existance of probabilitstic situations such as we find in quantum mechanics. This association of probability and uncertainty tends to suggest that, conversely, when 'certainty' exists probabilty (at least with a value between zero and one) is not involved. A further example suggesting a link between probability and knowledge, or rather lack of knowledge, is seen in the common practice of defining random sequences as those that make 'gambling systems', of a certain type, unprofitable. i.e. Random sequences are those for which the position selections of gambling systems return relative frequencies, in the limit, equal to the overall relative frequency of the sequence. Gambling systems can be considered as techniques of predicting the contents of sequence positions, and therefore as methods of embodying knowledge about sequences. But random sequences are defined as those that do not respond to the predictions of gambling systems, and are, therefore, as it were, those sequences about which our knowledge is at a minimum, or, if you like, those that maximise our uncertainty. Once again we see a case where the concept of probability is implicitly linked to uncertainty in knowledge.

If we accept that there is a link between probability and knowledge, we may well ask, is it that probability is an objective phenomenon which only as a by product has the effect of inducing uncertainty? Or is it that probability is an entirely subjective phenomenon that is synonomous with uncertainty? The view of probability I will be advancing proposes that probabilities can not exist without an observer knowing something, or rather not knowing it, and therefore it is not an intelligible notion without an observer in the picture. However, knowledge concerns, and is linked to the 'objective' world. Therefore, since probability concerns knowledge (or lack of it), we shall find that probability too has links with this 'objective' world.

3 A THEORY OF PROBABILITY

In the previous section recognition was taken of the fact that in some instances a particular outcome may have associated with it more than one 'probabilistic' magnitude 'P', depending on the various 'conditions' to which the outcome can be related. It was suggested that whether or not a tendered magnitude, 'P', is in fact a probability, depends in some way on an observer's knowledge of such relations, thus making probability relative to an observer. In this section I will enunciate a single principle defining probability which I hope will clarify and explain this dependence of probability on an observer's point of view.

• • • • •

Shortly after the year 1795 Laplace enunciated seven principles of probability that were to serve as the foundation for the mathematical development of probability theory in its classical form. One of these principles was a definition of probability, and in this connection we find in Laplace's paper, 'A Philosophical Essay on Probabilities', the following statement:

> 'The theory of chance consists in reducing the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favourable to the event whose probability is sought.'

He went on to define probability as:

> '. . . the ratio of the number of favourable cases to that of all the cases possible.'

If this classical definition is not actually circular, then it only just falls short of circularity . . . it is a hard job not to define the so-called 'equally possible' cases, that are enumerated to make up Laplace's ratio, as being 'equally probable'. However, Laplace also refers to the possible cases as 'equally undecided', and elsewhere he suggests that this indecision is subjective rather than objective in nature:

> 'Probability is relative, in part to this ignorance, in part to our knowledge. We know that of three or a greater number of events a single one ought to occur; but nothing induces us to believe that one of them will occur rather than the others. In this state of indecision it is impossible for us to announce their occurrence with certainty.'

By defining 'equally possible' cases as 'equally undecided', Laplace appears to opt for a subjectivist interpretation of probability. But that does not prevent circularity; probability, as defined by Laplace, itself appears to be a measure of subjective indecision expressed as a ratio calculated from two classes of 'equally undecided' cases. As he leaves us with no definition of 'equally undecided' we are therefore tempted to pick up his definition of probability and define 'equally undecided' as meaning 'equally probable'. Thus, probability is defined in terms of probability.

These two elements of subjectivism and circularity found in Laplace's definition are, I suggest, not so much faults, but a perception of two real features of probability. Firstly, consider the subjective component; if we think about Laplace's concept of 'all possible cases' then what can only be meant by this is a range of states, one of which may be the state of the system under consideration, but presumably because of the ignorance of the observer he doesn't know which state it is, although he does reckon to know that it is one of the said range. The elements of the class of 'all possible cases' attain their status of possibility simply because the observer does not know enough to exclude any one of these cases as being the actual state of the system. On this view an outcome or case does not possess 'possibility' by virtue of some property intrinsic to the objective world that makes it a 'possibility', but instead it possesses it through the observer's knowledge, or rather lack of knowledge; a state or outcome 'possesses possibility' not as a result of some potential for possibility in the objective world, but because an observer's ignorance prevents him from excluding it as the actual outcome or state of the system in question. Therefore it follows that the notions of 'possibility' and 'range of possibility' have a subjective component, as clearly what is 'possible' can vary from observer to observer according to how much they know. As an example think of the throw of a coin; it lands heads up and one observer looks at it—To this observer the outcome of tails for this particular throw is an impossibility because he knows it to be untrue, but for an observer who has not seen the result tails remains a possibility, even after the coin has landed with heads up.

The concept of possibility as an option not excluded by knowledge is relatively clear cut. But life is not as simple as this; in his definition of probability Laplace talks of 'equally possible' cases, as if, by implication, there can be degrees of possibility: Presumably, if cases can be 'equally possible' they can also be 'unequally possible'. It is at this point that the circularity of Laplace's definition becomes apparent, as it would seem that such inequality in the possibilities could only be expressed by referring to their probabilities, and therefore to talk of equal possibilities is really to mean equal probabilities. I am going to advance a proposition that explains the implicit circularity here as a result of systems always inevitably being subsystems nested within larger systems to any level of nesting.

Let me develop this proposition in an abstracted and general way as follows: Let us consider some outcome/property/state that could be realised within system 'n'. Let us assume that as far as we as observers are concerned, the system has available to it a total of T(n) possible cases, and that of this T(n) states, F(n) of them are favourable to the outcome/property/state in question. However, perhaps system n is a subsystem of system n + 1, and that within the context of system n + 1 the said outcome is actually favoured by F(n + 1) cases out of a total of T(n + 1) possible states of n + 1. The system n + 1 may, of course, itself be embedded within yet a larger system, n + 2, of which we know enough to say that the outcome in question, in this context, is favoured by F(n + 2) cases out of T(n + 2) possible states of n + 2. The resulting series of ratios found by dividing the favourable cases by the total possible cases for systems n to n + m need not be the same, consequently begging the question as to which ratio represents the probability of the outcome/property. Another and equivalent way of looking at this is to realise that each of the T(n) possible states of system n, when viewed within the context of n + 1 are themselves outcomes that, depending on our knowledge, may be favoured by a certain fraction of possible states of n + 1, and therefore on Laplace's definition each of the T(n) possible states will have their own 'probabilities' within n + 1. Likewise, in turn, the possible states of system n + 1 will have, depending on our knowledge of the possible states of n + 2, associated with them Laplacian 'probabilities'. The problem that Laplace has given us here is that the probability within, say, system n can be calculated as F(n)/T(n) only if the T(n) cases are 'equally possible', or as I am interpreting this to mean, 'equally probable', which is not necessarily true because the nesting of system n within n + 1 may mean that its possible states are not equally probable. Thus, each state of n is, as it were, 'weighted' by its probability derived from its context in n + 1. This suggests, therefore, that the probability of an outcome in n is in fact a ratio of the summations of the relevant probabilities. i.e.:

$$\text{PROB(outcome in n)} = \frac{\text{Sum of the probabilities of the favourable states of n}}{\text{Sum of the probabilities of all possible states of n}}$$

$$= \sum P_f / \sum P_t \qquad (6.1)$$

Where the f are the favourable states and the t are the possible states, and where the P_f and the P_t are their respective probabilities calculated in system $n+1$.

If we consider this to be a definition of the probability of an outcome in n, then it obviously leaves us with a regression problem, whereby the probability of the outcome can not be determined until we determine the probability of system n's possible states within context $n+1$, which in turn can not be done until the probabilities of $n+1$'s possible states are determined within context $n+2$, and so on. However, if we accept the definition for the moment then we can see that it is actually a generalisation of Laplace's definition; the latter being a special case where all P_f and P_t are equal, giving rise to the consequence:

$$\text{PROB(outcome in n)} = F(n)/T(n) \qquad (6.2)$$

The definition of probability given in equation (6.1) defines probability in terms of probability, but it is not so much circular as it is recursive. The question is, when does the recursion stop?; this is something Laplace did not make clear. Successive probabilities are indeterminate because they derive from probabilities relating to the next level up which themselves are indeterminate. However, the recursion would be prevented from indefinite regression and the problem solved, if at some level, say $n+m$, the probability of an outcome of this level was simply defined, as Laplace suggested, as the ratio of the number of favourable states to the number of all possible states that $n+m$ could as far as we know take up. If this was the case then probabilities at successively lower levels would be inferable.

In fact there is one system that naturally qualifies for an application of Laplace's original definition; that system is the universal system, the system that encompasses all things. For the universal system, u, there is no greater system, $u+1$, of which it is part. Therefore, the possible states of the universe (which attain their possibility because ignorance prevents their elimination), can not be seen in a higher level context that gives them probability; the probability of these states can not be expressed using an equation similar to equation (6.1) because this equation depends on the existence of a higher level system, $u+1$. Thus, the exceptional nature of the universal system suggests we postulate that its possible states have equal probability, implying that in principle, the probability of any outcome within this system could be expressed using the Laplacian ratio of equation (6.2), and where $n=u$.

If we accept the equal probability of the universal states, then this allows us to identify circumstances under which the states of a subsystem are also equally probable, and therefore means that we can express the probability of outcomes within the subsystem using equation (6.2). These circumstances exist whenever the possible states of a subsystem are favoured by equal numbers of universal states; since the latter are equally probable then so, by implication, will be the subsystems states, thus allowing the application of equation (6.2) to its outcomes. We can identify more than one type of subjectively relative circumstance that gives equal probability to the possible states of a subsystem. For example:

(i) If the knowledge of the observer is such that the possible state of the surrounding universe is irrelevant to the state of the subsystem; i.e. Relative to the observer, all combinations that can be formed from the possible state of the subsystem and the possible states of the rest of the universe are themselves possible. This will mean that each of the subsystem's possible states are favoured by equal numbers of the possible universal states because each possible state of the subsystem could, as far as the observer is concerned, coexist with any of the elements of the same class of possible states of the surroundings.

(ii) If the knowledge of the observer is such that the possible states of the surrounding universe relevant to the state of the subsystem are of such a small percentage that they can be neglected; i.e. States of the surrounding universe that are known not to be able to coexist with certain possible states of the subsystem are few enough to be neglected. This will mean that the point made in (i) above will approximately apply.

(iii) If the knowledge of the observer is such that the number of possible states of the surrounding universe relevant to the state of the subsystem can not be neglected, but the numbers of universal states favouring each subsystem state are even; i.e. Each possible subsystem state can, relative to the observer, coexist with the same number of possible states of the surrounding universe, although in each case these numbers may enumerate a different class of states.

In each of the above observer relative circumstances the possible states of the subsystem in question are equally probable thus allowing us to apply Laplace's original definition (equation (6.2)) to the probability of outcomes within this subsystem.

Comments: The above definition of probability uses a concept of 'possibility' that has a subjective component, as clearly what is possible to one observer through his ignorance, may by another observer be known to be impossible. Consequently, probability itself may be observer dependent. In the next section some applications of this notion of probability will be explored. However, one point I would like to stress is that probability is calculated from classes of possibility the observer **reckons** to hold for the system under consideration. This means that even though the actual state of the system is not one of the observer reckoned possibilities (i.e. he falsely believes the true state of the system to be impossible), the calculated probability with respect to this observer still applies because it is a ratio derived within the context of the knowledge the observer reckons to hold. In spite of the subjective account of probability I have given, this does not mean that probability is not linked to objective conditions. An observer's view of the range of possibilities is connected with the objective world because the implied assertion that a system must have a state within certain limits, if this assertion has extra-logical content, means that it can be true or false, depending on the actual state of the system. Consequently, just as one finds the probability of an outcome is a function of the relativity of the knowledge of limits, one also finds that it is linked to the objective content of this knowledge, and in some cases its value may correspond to objective parameters: As an illustration consider a set of items, each of which may or may not possess a property of a certain type. Imagine that an observer asserts that the property is distributed over the set of items with a relative frequency of R%. This assertion implies that the system of items is limited to those distributions of properties that are consistent with the value of R%. It is easy to show that of these distributions the fraction of cases favourable to an item possessing the property in question, is equal to R%. Therefore, with respect to the observer, R% will be numerically equal to the probability of an item possessing the property, if this observer's information is such that for him the cases consistent with this value are equally possible. We see here an example where the probability of an outcome/property is linked to an observer's objective description of a system to such an extent that its value is actually found in that description. The probability arises out of the description that the system has a relative frequency of R% because this description is incomplete; this incompleteness exists because the value R% admits a range of distributions consistent with it. Thus, because this range also represents a range of possibilities for the observer then we can see that a probability ratio will arise out of this range. Probabilities, it seems, are linked with incomplete descriptions of a system (i.e. descriptions that do not assert the exact state of the system but tie it down to a range of states), but they are only implicit in a description if it represents an observer's full knowledge of the system when, therefore, the implied range of states is also a range of possibility for the observer concerned. When this subjective condition is fulfilled it is then that the numerical magnitudes of the probabilities will be implied by the description, and may, as we have seen above, correspond to values of objective parameters that the description asserts.

4 APPLICATIONS

The theory of probability proposed in section (3) is now going to be applied to throw light on a variety of connections.

(a) *Random/disordered sequences*: According to the account of probability I am proposing probability is linked to the objective world in that its value is implicit in descriptions about this world if the relevant observer is unable to provide more informative descriptions. This is a rather general explanation of the connection between the objective world and probability which does not explain the peculiar fact that certain situations, such as irregular or disordered sequences, are far more closlely associated with probability than many others. The association of random/disordered/irregular sequences with probability is so strong that it has induced philosophers of the past to define probability in terms of the relative frequency of these kinds of sequence. This association must therefore be acounted for.

Let me start by noting that randomness has an important feature that on one occasion at least has been considered as its defining characteristic. This characteristic can be expressed as the 'independence' of the results of particular outcomes in the sequence of each other. 'Independence' in this context is understood to have a configurational meaning: That is, independence is said to exist, if over any reference class of trials consisting of all those trials immediately following a particular configurational type, a relative frequency is returned that is equal to the RF, R%, of the sequence as a whole. 'Independence', or 'freedom' as he called it, was used by Karl Popper in his 1934 'Logic of Scientific Discovery' as a definition of randomness. This feature of random sequences is a clear cut and well defined concept, which, philosophically speaking, can be identified with relative ease as an objective characteristic with its basis in frequency enumerations. This measurable characteristic, when it exists, may favour the existance of subjective probability for reasons that I shall now explain.

The overall relative frequency of a sequence, R%, is an unambiguous and objective characteristic of a real sequence. But this objective quantity, even for a random sequence does not in itself entail a probability of P = R%. However, if an observer's only information concerning a particular trial is that it is part of a sequence with an overall RF of R%, then R% is numerically equal to the probability of the outcome of the trial for the observer concerned. This follows if each configuration of the total sequence consistent with the value R% is equally probable, because then it can be shown that the ratio of the number of configurations favourable to the trial possessing the outcome in question to that of the total possible configurations, is equal to R%. Thus, under these subjective circumstances R% is numerically equal to the probability of the outcome for the observer concerned. Now, the interesting thing about a sequence that is random is that an observer is unable to alter this subjective estimate of the probability on the basis of any knowledge he may possess about preceding local configurations. This is because the 'subsystems' formed by selecting classes of trials according to the form of immediately preceding local configurational types all have relative frequencies of R%. Thus, for an observer who only knows that the sequence is random with an overall RF of R%, the fraction of cases favourable to an outcome in each of these subsystems is always R%; i.e. Every outcome will have a probability of R%, regardless of what local confiuration is known to procede it. Knowledge of randomness and local configurations is not sufficient information to differentiate the probabilities of each outcome, but instead this informational basis will mean that the probabilities remain constant, independent of sequence position. Consequently, because our own point of view means that we often find ourselves in a position where we believe we know that a sequence is random and that it possesses a certain RF, but do not know much else, it is not surprising that random sequences are strongly associated with single, unchanging probabilities.

The above account of the relation between probability and randomness contains some mathematical and philosophical issues that need attention. Firstly, mathematically speaking, probability as I have defined it requires a more rigorous treatment in its relation to randomness than I have suggested above, by allowing for the complexities introduced by the sequence concept and the given randomness of the sequence. If we are considering an unsequenced set of items (or trials) with a distribution of R%, the fraction of all configurations of properties (or outcomes) consistent with R% that are favourable to a particular item possessing a property, is numerically equal to R%. Thus, if for the observer concerned the configurations consistent with R% are equally probable, then it follows that R% is numerically equal to the probability of the property. However, for a sequence of items (or trials) the logical constraints introduced by sequencing and the stipulation of randomness, means that in a subsystem of items selected using a local configurational type, any observer would have enough logically implicit information to eliminate a percentage of the configurations of properties in this subsystem. Thus, not all of the configurations of properties within the subsystem which are consistent with R% are also consistent with these extra constraints, and therefore R% does not necessarily give the fraction of configurations favourable to an item (or trial) possessing a property (or outcome). In a more rigorous treatment we would have to show that this 'mark down' in the configurations consistent with R% is not sufficient to effect the value of the probability. At this point it is convenient to introduce another issue of more philosophical character that naturally arises out of this consideration. This is the question as to whether logically implicit information ought to be included in a calculation of probabilty, even though the observer concerned does not know the logical implications of what he does know. That is, is probability equal to the ratio of logically favourable cases to that of all logically possible cases, or is it equal to a ratio of cases that are only directly known by an observer to be possible? I am inclined to define probability in the latter sense, as it gives it a wider application and would mean that the notion of probability could be invoked where complex and difficult calculations are effectively a barrier to knowledge.

(b) *Ordered sequences*: Although random sequences may favour the existence of relative probabilities, this does not mean that ordered sequences are excluded from returning probabilities; In general, probabilities arise relative to an observer if a system, like a sequence, is at most known to have a state in a range of possible configurations. This range may include both ordered and disordered configurations. The result is that both ordered and disordered configurations may have comparable probabilities. This consequence enables us to explain the paradoxes raised in section (1) concerning the association of probabilities with regular sequences. The high, and in fact overwhelming probability associated with disorder as a condition represented by a class of sequences, arises because this condition tends to be shown by a large majority of possible configurations, thus giving a large majority of cases favourable to it, and therefore, in turn giving it a high probability.

(c) *Statistical Validity*: Consider a system with N equally possible states. Let an outcome, O_i be realised by s_i of these N states. Let us require that $s_i + \delta_i = N$, where δ_i is small, so that the probability of O_i, represented by $P(O_i)$, is almost one. Let us now imagine that we have outcomes O_1 to O_n, all of which have a similarly high probability for the system concerned. This means that the quantities represented by the δ_i are all small compared to N. A particular δ_i is, in fact, the number of cases favourable to outcome O_i not occurring. Thus, if we take the union of the two sets enumerated by δ_i and δ_{i+1}, this unified set represents the cases favourable to either of O_i or O_{i+1} not occurring. Therefore, the union of all sets enumerated by the δ_i represents the cases favourable to an outcome where at least one of the O_i does not occur, or in other words, the cases favourable to the conjunction of the O_i being false. Now, there is no reason that makes it impossible for this unified set to have a size equal to N, thus implying that it is possible for there to be no cases favourable to the conjunction of the O_i. This logically admissible conclusion represents the case when the conjunction of the O_i has a zero probability. Such a conclusion is allowed by the concept of probability I have proposed, even though each individual outcome, O_i, may have a probability close to one.

The obvious application of the above consideration is to the issue raised in section (1), where it was found that there was a very high probability of each and every sequence reference class returning statistically valid frequencies, and yet it is not possible for all reference classes to be statistically valid at the same time.

(d) *The Coin*: Imagine that before us we have a coin on a table and that the coin is covered by an upturned cup, so that we can not see and do not know which face is uppermost. What is the probability of the coin being tails up? If we know no more, then from our point of view the system has two possible states (heads up and tails up), one of which favours the said outcome. Thus, the probability of the coin being tails up is $\frac{1}{2}$. Now, this calculation seems too easy, too trivial, and too good to be true. It may even seem wrong, because if the coin has been tossed on to the table there is no logically obliging reason why tails should have a probability of $\frac{1}{2}$, as perhaps the coin is biased. But this depends entirely on our knowledge of the context in which the outcome is viewed; in the descripton of the system of the coin on the table, no information is given about constraints linking its state (tails up or tails down) with the state of any other system. Thus, if this description is all we know, then for each of the possible states of any other system, the coin could possibly be tails up or tails down. The implication of this, relative to our position of ignorance, is that each state of the coin is favoured by the same number of equally possible states of the universe. Therefore, given this degree of knowledge of the coins context, we will return an equal probability for tails up and tails down of value of $\frac{1}{2}$.

Alternatively, we as observers may know a little bit more about the coin's relation to its context than stated in the initial description. We might know, for example, that the coin was thrown on to the table, and is part of, or hypothetically could be part of, a series of throws that has a relative frequency of tails of R%. Given knowledge of this related context, hypothetical though it may be, means that for us tails up and tails down are no longer necessarily equally possible, as they are outcomes in a larger system of which our knowledge is such that their probability ratios in this context are R% and 1 − R% respectively.

(e) *The Three Coin System* (see section 2): The problem that this system poses is embodied in the following question: What is the probability of getting heads?; is it P1, P2, or $\frac{1}{2}(P1+P2)$? To answer this question I have to start by assuming that the numerical magnitudes of P1 and P2 correspond to the relative frequency of heads over a sequence of real or hypothetical throws of the respective coins (the sequences to which these relative frequencies apply are usually considered as unlimited). In making this assumption I am recognising that a sequence of throws, actual or hypothetical, will have an unambiguous value of relative frequency unless we are to entertain the nebulous idea of a sequence somehow possessing more than one RF value at a time. Thus, P1, P2, and $\frac{1}{2}(P1+P2)$ are considered to correspond to objective quantities, but they can not be directly equated with probabilities. In this particular case probabilities will only arise as a function of how much is known about the conditions relevant to a trial of the system and whether or not it can be related to particular coins and therefore to a particular sequential context. For example, if all the observer knows is that the throw of a coin, actual or hypothetical, is a product of the three coin system (i.e. he does not know whether it is of the biased or unbiased coin), then the probability of heads is unambiguously equal to $\frac{1}{2}(P1+P2)$. This follows because given the information the observer has, the outcome maps to a system where the fraction of cases favourable to it being heads out of all cases the observer reckons as possible is equal to $\frac{1}{2}(P1+P2)$. (We assume the cases here are equally probable.) This situation changes as soon as the observer is aware of which coin (i.e. whether the biased coin or unbiased coin) was involved in the trial. In this case the observer is able to map the outcome to a particular coin which provides a sub-context giving the fraction of equally possible cases favourable to heads of either P1 or P2, depending on which coin is known to have been used. It is worth noting the change in the status of the magnitude $\frac{1}{2}(P1+P2)$ once the observer is aware of the new mapping; it no longer qualifies as the probability of heads since it no longer measures the fraction of favourable cases, but from an objective point of view it is still a valid measure of a characteristic of the three coin system.

(f) *A Deck of Cards*: Imagine that we have on a table in front of us what we know to be a half pack of playing cards (i.e. twenty-six cards). Assuming that we do not know how the cards are sequenced in the pack, what is the probability of the top card being an ace of diamonds? The answer to this question may seem to depend on whether or not the half pack actually contains the ace of diamonds; if it doesn't then the probability is zero; if it does then the probabilty is 1/26. Is this correct? No, it's wrong. In the initial statement of the problem we are only given the information that the pile of cards is half a pack of playing cards. We are not told whether or not the pack contains the ace of diamonds. Consequently, as far as we the observers are concerned, the top card could possibly be any one of fifty-two cards, implying that the fraction of cases favourable to it being an ace of diamonds is 1/52. Thus, assuming the possible cases are equally probable, the probability of ace of diamonds is 1/52. Whether or not the half pack contains the ace of diamonds does not effect this result. (Although the probability does change if we are given information on the whereabouts of the ace of diamonds.)

Now, this example illustrates how different the objective characteristics of a system may be when compared with an observer reckoned probability. If the half pack does not actually contain the ace of diamonds then objectively speaking it is impossible for it to return an ace of diamonds as the top card, even though the probability of this outcome is 1/52, and therefore subjectively speaking possible. Alternatively, we might imagine that the half pack contains the ace of diamonds and is part of a system consisting of a potentially unlimited sequence of shufflings and of placing on the table. This system may well yield a relative frequency of ace of diamonds on the top of the pack very different from 1/52, but still, given the information in the initial statement of the problem, the probability remains as 1/52.

. . . . .

There is one final point to discuss in connection with this latter application. Although a series of shuffles and placings was mentioned, nothing was said in the initial statement of repeated shufflings or even whether the pack was shuffled at all. In fact there need be no dynamics in this system for it to return a relative probability. To arrive at a relative probability of 1/52 the system does not necessarily have to be capable, as Kolomogorov would have it, of an infinite series of repetitions. The cards could have glued together since the day of creation for all we know, but as far as our information is concerned, the top card being an ace of diamonds is still favoured by 1/52th of the cases that we can not rule out as impossible. A similar point can be made in connection with the application (d) involving the coin on the table; there is no reason why the coin should have been placed by dynamic means, or be part of a series of tosses of infinite extension—the reality of the coin on the table may be all there is to it.

We see that these two applications are capable of returning relative probabilities even though they are static systems with no moving parts, and therefore possessing nothing like an objective 'propensity'. Similar considerations arise more generally in connection with 'historical' and 'static' distributions. Here we have no dynamics, no moving parts, and no balance of potential capable of producing anything new, and yet it is perfectly natural, and in line with my proposal to use (subjective) probabilities when referring to the outcomes/properties in these distributions. Systems such as this have but one timeless state, which, however, if viewed from the point of view of an observer with limited knowledge, may give every appearance of producing something new and surprising at every outcome, particularly if the distribution is disordered. The notion of probability I am advancing is such that it can be applied to what has been, just as well to what is, what will be, or, what 'would be'. According to this view, time, dynamics, and process are irrelevant to calculations of probability.

Norwich, England

I REFERENCES

Gillies, D. A. [1973]: 'An Objective Theory of Probability.' Methuen 1973.
Laplace, P. [c 1795]: 'A Philosophical Essay on Probabilities.' Dover Publications. 1952.
Mellor, D. H. [1971]: 'The Matter of Chance.' Cambridge University Press. 1971.
Popper, K. R. [1934]: 'The Logic of Scientific Discovery' Hutchinson, Sixth Impression 1972.

Brit. J. Phil. Sci. 39 (1988), 183–192 *Printed in Great Britain*

Absolutism and Relationism in Space and Time: A False Dichotomy*

IAN HINCKFUSS

ABSTRACT

The traditional absolutist-relationist controversy about space and time conflates four distinct issues: existence, abstraction, relationality and relativity. Terms which are relational, relative or abstract may denote items which possess contingent properties. Possession of such properties, including topological and geometrical properties, is therefore no indication of logical type. To fail to recognise the possibility of spaces, times and space-times of various logical types is to risk conflating two distinct ontological issues: a metaphysical issue concerning the existence of abstract objects and a question of physics concerning the existence of causally efficacious sub-strata which may or may not be needed to explain the contingent properties of the abstract objects.

1 *Introduction*
2 *A Theory of Abstract Non-Relational Time*
3 *An Analogy—Abstract Santa and Person Santa*
4 *Conclusions*

I INTRODUCTION

There are dozens of logically distinct, though not necessarily independent issues involved in the absolutist-relationist controversy about time and space. The late John Mackie in his [1981] isolated fourteen of them—but there are more.

One may adopt what I shall call a resolutionary stance by embracing the relationist position on some issues and the absolutist position on others. More interesting resolutionary approaches involve the disambiguation of some of the issues or the discovery of false presuppositions common to both the relationist and the absolutist.

One presupposition of the debate seems to be that there is one and only one concept of 'time' and one concept of 'space' that we are exploring. In this paper six different categories into which notions of 'time' or 'space' may fall are presented. There is no question of any one of these being right and the others being wrong. The question of whether 'time' or 'space' in ordinary language falls into one or another is irrelevant. The fact is that our language is sufficiently rich if not adaptable to accommodate all these concepts. What is important is that we know what we are talking about in any one case.

---

* Thanks are due to an anonymous referee whose comments enabled me to improve an earlier version of this essay.

Much recent philosophical work on time and space seems to presuppose that contingent properties of time and space, particularly those which may feature in explanations of physical events, constitute *prima facie* evidence for the absolute nature of time or space. Graham Nerlich has suggested that this applies in particular to certain topological and geometrical properties of space and space-time. Such supposedly *prima facie* evidence for absolutism is then supposed to present a challenge to the relationist to produce a 'reduction' of the statement in question with a view to quashing the *prima facie* case.[1]

In the conclusion of *The Existence of Space and Time*, I argued that, given an appropriately broad notion of 'reduction', such a reduction was always possible. This argument has been criticised by Chris Mortensen [1977] and Nerlich [1979], but these papers do not question the presuppositions of the debate.

In this paper I wish to argue that the presuppositions of the debate are incorrect. The fact that time, space or space-time may contingently exhibit some property or other does not constitute even *prima facie* evidence for the absolute nature of time, space, or space-time, in any of the multiple senses of 'absolute'. I shall also show that two quite distinct ontological issues are being conflated in the debate.

I shall begin with a sketch of a theory according to which temporal reference in ordinary discourse is abstract but non-relational. I shall then compare the objects of this theory with relational abstractions which I call 'Abstract Santa' and 'Relative Abstract Santa' to demonstrate how such abstractions and relational and relative terms can feature in causal explanation without themselves being regarded as referring to anything causal. It does not matter, for my purposes, whether the theory of temporal reference is true or even plausible. I happen to believe it is, but that is irrelevant for the purpose of this exercise for which it matters only that the theory makes sense.

2 A THEORY OF ABSTRACT, NON-RELATIONAL TIME

Consider any utterance of:

A. John kissed Mary last Thursday.

or, in other words:

B. Last Thursday was a time when John kissed Mary.

---

[1] Much of Hinckfuss [1975] is based on such presuppositions. Nerlich's arguments appear in his [1973], [1976] and [1979]. Sklar's reductionist reply [1974] to Nerlich's [1973] also seems to be in line with these philosophical presuppositions.

The words 'last Thursday' in A, modify the proposition expressed by 'John kissed Mary' to a new proposition which entails but is not entailed by 'John kissed Mary'. The new proposition, expressed by A, refers to a time during which John kissed Mary, namely, last Thursday. In sentence B, which is synonymous with A, 'Last Thursday' occurs, not as a propositional modifier, but as the subject of a sentence, as a noun phrase referring to a time, a time when John kissed Mary.

Compare A and B with C and D:

C. John kissed Mary tenderly.
D. Tenderly was the way in which John kissed Mary.

In C, 'tenderly' acts as a propositional modifier. In the synonymous D, 'Tenderly' is transformed into a noun referring to a way, a way in which John kissed Mary.

Ways, manners and modes are abstractions corresponding to propositional and predicate modifiers. Analogously, it might seem reasonable to conjecture that times are abstractions corresponding to temporal propositional modifiers like 'last Thursday' in A. The idea is not new. The seeds were sown in Locke's *Essay Concerning Human Understanding*, Book II, Chapters 13 and 14, where he suggests that we can get new ideas by modifying ones we already have.

Those who wish to take a more substantive, non-abstract view of times may insist here, that we must not allow ourselves to be misled by grammatical analogies. They could point out, for example, that A and B could be compared with E and F instead:

E. Mary rode Cedarlegs.
F. Cedarlegs was the mount that Mary rode.

They might even allow that we treat 'Cedarlegs' in E as a modifier of the proposition expressed by 'Mary rode', just as long as we do not allow that to lead us to believe that Cedarlegs, or any mount for that matter, is less substantial than any mount has to be. No mount that Mary rode was a mere abstraction.

Very well, the grammatical analogies do not prove anything, but at least they show us what sort of abstraction, if any, times may reasonably be thought to be. Though abstractions, they would not correspond to monadic or polyadic predicates. Rather, they would be abstractions corresponding to certain sorts of propositional modifiers.

In *Some Uses of Type Theory in the Analysis of Language* [1974], Malcolm Rennie[2] developed Church's formulation of the simple theory of types for application in the theory of predicate modifiers among other linguistic items. A hierarchy of types is recursively defined with a basis of *individuals* (type i), *truth values* (type o) and what he and many others have called *worlds* or *possible worlds* (or what I prefer to call possibilities) (type K). The inductive clause in the definition is that if α and β are types, then the functions from items of type β to items of type α constitute a type. The type of functions from items of type β to items of type α is designated $(\alpha\beta)$.

---

[2] I take this opportunity to pay tribute to Malcolm Rennie—former colleague, mentor and friend. His mental abilities and his logical and philosophical acumen were outstanding. With those who were not so intellectually endowed he was extremely patient, helpful and encouraging. He is sadly missed.

Thus if propositions are functions from worlds (type K) to truth values (type o), then propositions are of type (oK). Functions from propositions to propositions (such as negation, necessity, and last Thursday (if our theory of temporal abstraction were correct) are of type ((oK)(oK)) or $(oK)^2$ for short. Properties of items of type α yield propositions about those items and hence are of type $((oK)\alpha)$. Dyadic relations between items of type α are of type $(((oK)\alpha)\alpha)$ and n-adic relations between items of type α are of type $(\ldots((oK)\underbrace{\alpha)\ldots\alpha)}_{\text{n occurrences}}$.

Hence properties of individuals and n-adic relations between them would be of type ((oK)i) and (. . . ((oK)i). . . i) respectively, and hence would be of the wrong type, within the theory under consideration, to feature as times like last Thursday which would be of type $(oK)^2$. So relationists who, like Leibniz, would think of times as relations between individuals in the form of events or otherwise are not thinking of times as abstractions from temporal modifiers. Neither are absolutists who believe that last Thursday is of type i along with tables, atoms, cats and dishwashing machines. Both relationists and absolutists have traditionally held an impoverished view of abstraction. If one allows that all substances, material objects, processes and events are individuals, and that all non-individuals are abstractions, then times like last Thursday, granted that they are of type $(oK)^2$, are certainly abstractions.

Assume for the moment that times are of type $(oK)^2$. That does not differentiate them from many other items of that type, including places. Anything that has so far been said of last Thursday could have been said of a place—in Paris, for example. Being of type $(oK)^2$ would be a necessary condition for being a time, but it would not be sufficient.

On any account, times, like material objects and places, must be able to enter into mereological relationships with one another. Thus, last Thursday was part of last week. It was also the fusion of the set of all those hour-long periods which occurred last Thursday.

Material objects, places and times may all be extended. Material objects are usually thought of as four-dimensional (three spatial and one temporal). A place may have from zero to three dimensions depending on whether it is a point, line, surface or volume. A time is usually thought of as a one-dimensional thing unless it is of zero duration. The logical possibilities of material objects having more than four dimensions, times having more than one dimension and places having more than three dimensions need not be excluded. In any case, the usual mereological axioms will apply.

Take any time such as last Thursday and consider the set of all the times of which it is a part. Form the fusion of this set and let us call such a fusion a Time (with a capital T). The development of the theory of relativity has raised the possibility of there being more than one Time, indeed, of there being an infinity of them—one for each inertial frame of reference. Similarly, Space (with a capital S) has blossomed into an infinity of Spaces with Relativity Theory.

Times (with capital T) and Spaces are usually considered *dense*. That is to say, for any two non-intersecting times within a Time, there is at least one time that is between the other two. The betweenness relation applicable between the times of a Time determines the topology of the Time. Thus if it were true that for any three times t_1, t_2 and t_3, if t_2 were between t_1 and t_3, t_3 was not between t_2 and t_1, then time would possess no closed loops.

Enough has been said to establish the following. Some of the properties of Times seem to be a function merely of our conceptualisations, for example, (let it be allowed) the density of Times. (This is not to say that we are stuck with dense Time in the sense that we may never be willing and able to use a notion otherwise like that of 'Time', but to which the axiom of density does not apply of logical necessity.) On the other hand, a Time may exhibit properties contingently, for example, the topological property of having loops or otherwise.

But how would Time, if it were an abstract construction out of abstract times, come to have such contingent properties? What are these contigent properties contingent upon? To say 'the way the world is' would be empty if by 'the world' we meant 'everything that is the case'; for part of what would be the case is that Time would have the contingent property to be explained. On the other hand, if by 'the way the world is' we mean 'the way the *furniture* of the world is', that is, 'the way the *individuals* of the world are', that is, 'the way the real world is' where what is real is to be contrasted with the abstract artificialities of our conceptualisations, then what we have is a statement of reductionism—one of the tenets of what is vaguely known as relationalism.

The sort of reductionism just outlined is very weak. Indeed it is consistent with this sort of reductionism that there exists one or more substantial Process-Times as individuals of the world, such that these and perhaps even only these are needed to explain the contingent properties of any Abstract Times. A stronger reductionism is to be had by the addition of the proposition that there is no Process-Time among the individuals of the (actual) world, with the implication that the explanations of the contingent properties of Abstract Times are to be generated in terms of individuals other than Process-Times.

Similar statements apply *mutatis mutandis* for Abstract Space and Substance Space. The idea of there being more than one sort of space is not new. Newton himself recognised the existence of observer-relative abstract spaces as well as the more substantial absolute space which he thought was necessary to explain the phenomenon of inertia.

3 AN ANALOGY—ABSTRACT SANTA *VS* PERSON SANTA

To illustrate the situation with a partly analogous case, consider the widespread belief among young children that a person called 'Santa' gives gifts to people at Christmas. Let us re-name Santa with the name '*Person-Santa*' to differentiate him from *Abstract-Santa* which is hereby defined as the relation of giving at Christmas. Abstract-Santa is exemplified by many triples such as, for example, that consisting of Mary, Alice and the box of lollies Mary gave Alice last Christmas.

Let us say that something is *Christmas-generous* if it is a relation such that the number of Christmas gifts in its field is equal or greater than the number of people in its field, or is a person who gives at least as many gifts at Christmas as there are people in the world, and let us say that something is *Christmas-lousy* otherwise. Then the existence of a Christmas-generous Person-Santa will ensure that Abstract-Santa is also Christmas-generous.

The converse, however, does not hold. Abstract-Santa is Christmas-generous if and only if there are at least as many Christmas gifts as people. But this can be so without there being any person (let alone Person-Santa) who would be willing and able to distribute that many gifts at Christmas.

However, if there is no Person-Santa to explain the Christmas-generosity of Abstract-Santa, the Christmas-generosity of Abstract-Santa is still to be explained by such gift-giving as there is of ordinary mortals, Christmas-lousy though they all may be.

Note that if anything is a relational concept of Santa-hood, Abstract-Santa would have to be that. For Abstract-Santa, by definition, is a relation. You cannot get more relational than that. Yet nominalism aside, Abstract-Santa exists and is unique. Existence and uniqueness are often taken to be signs of an absolute entity, but where 'absolute' is meant to be contrasted with 'relational', they are neither necessary nor sufficient conditions for being absolute. For example, the space of which Leibniz writes in his letters to Samuel Clarke seems to be the set of all spatial relationships between material objects. Again ignoring nominalism, such a space exists and is unique; yet Leibniz's beliefs have traditionally been dubbed 'relationist' not 'absolutist'.

There is, of course, a sense of 'absolute' meaning 'non-relative', where an attribution of ø (say) to something is said to be absolute as opposed to relative-to-some-other-thing. Thus what counts as currency is relative to countries and what counted for Newton as a relative space was relative to an observer or frame of reference. If ø is a relative description of x, then 'x is ø' is incomplete; that is, to obtain a truth value for an utterance of 'x is ø', one must be able to supplement the information with the answer to the question 'With respect to what?'. The term 'Abstract-Santa', by contrast, is quite absolute in this sense. To describe something as Abstract-Santa yields a truth or falsity in itself. The sentence would be complete. No further reference to some item to which Abstract-Santa is related is needed to yield truth value.

We could, of course, easily develop a relative notion, call it 'Relative-Abstract-Santa' or 'RAS' for short, which was relative to countries. The RAS for some country, C, would be the Abstract-Santa relation restricted in its field to the residents of C. Then if someone claimed that RAS was Christmas-lousy, we would have to know what country they had in mind before their statement could be given truth value. With all due respect to Muslim countries, I guess their RAS would be Christmas-lousy, whereas the RAS of most Christian or post-Christian countries would be Christmas-generous.

'RAS' would be correctly used only in referring to certain abstract relations. But not all relative expressions are like that. Some are used correctly only if they describe people or other individuals, e.g. 'father', 'closest friend', 'nearest material object' and other such functional expressions. Such terms are not relational. They do not refer (directly at least) to relations between objects or anything else. Neither fathers, closest friends nor nearest material objects are abstract relations. They are only members of the fields of relations.

Neither Abstract-Santa nor any RAS is physical or is any sort of individual in the sense used here. Though statements describing Abstract-Santa can be contingent, Abstract-Santa is not causally responsible for the existence of the situations corresponding to true contingent statements about Abstract-Santa. Rather, it is the people in the field of the Abstract-Santa relation who are causally responsible for the situation. That is why we think of these people as part of the world—part of reality. The situation described by 'Abstract-Santa is generous' can be causally efficacious, but Abstract-Santa itself is not causally efficacious—it is a mere abstraction.

Correspondingly, if it is shown that some situation is causally efficacious and hence physically explanatory and that the situation is correctly described by 'Time (or Space) is ø' for some term ø, it does not follow that the Time (or Space) we are talking about is causally efficacious in itself[3]—it could be something quite abstract.

4 CONCLUSIONS

There are at least four distinctions that tend to be overlaid in the absolute-relational time and space controversy:

---

[3] Nerlich [1979] allows that Space may not be casually efficacious in the sense that something is causally efficacious only if 'it exerts a force on something, if it changes something's momentum or exchanges energy with it, or if it is involved in action-reaction pairs'. (p. 71) The claim I make here, however, is stronger than that. I am using a broader sense of causal efficacy wherein I allow that events, situations and processes could have causal efficacy. None of these satisfy Nerlich's narrower sense of causal efficacy.

terms with a null extension	*vs*	terms with non-null extension
terms which denote abstractions	*vs*	terms which denote non-abstract, concrete or substantial objects
relational terms	*vs*	non-relational terms
relative terms	*vs*	non-relative terms

Forgetting about existence for the moment, and allowing that all relational terms are terms denoting abstractions, these yield six possibilities as follows:

(i) relational, relative terms denoting abstractions
(Examples: RAS, Leibnitz space relativised to finite regions)
(ii) relational, non-relative terms denoting abstractions
(Examples: Abstract-Santa, Leibniz Space.)
(iii) non-relational, relative terms denoting abstractions
(Examples: Abstract-Times, Abstract-Spaces in Relativity Theory, Newton's Relative Spaces,[4]
last Thursday (relative to standards such as Eastern Standard Time).)
(iv) non-relational, non-relative terms denoting abstractions
(Examples: Abstract-Time in a Newtonian universe, Abstract-Space-Time in a Minkowskian universe, cardinal numbers.)
(v) non-relational, relative terms denoting substantive individuals
(Examples: the Prime Minister, the nearest tree)
(vi) non-relational, non-relative terms denoting substantive individuals
(Examples: Chris Mortensen, The Eiffel Tower, Absolute Space and Absolute Time in a Newtonian universe, Substantial Space-Time in a Wheeler universe.)

In all six categories, there are terms that refer to things that may be bearers of contingent properties. Moreover, in all six categories, there are terms which refer to things which can be the bearers of contingent geometrical and topological properties.

Further, note that relative terms may pick out items that have properties that are absolute in the sense that these properties do not vary from one of the items to another. Thus, all RAS's may be Christmas-lousy and all Abstract-Spaces in a relativistic universe may be Euclidean. Statements which expressed such propositions could be causally explanatory insofar as they would describe situations (of type i) which cause other situations or phenomena. Where such statements were geometrical or topological, the explanation could be described as geometrical or topological also.

Relative terms may pick out items which have absolute properties in the above sense. Also absolute (in the sense of non-relative) terms can pick out items describable with relative terms.

---

[4] In a previous draft of this essay I classified Newton's Relative Spaces as relational. Steven Savitt suggested correctly, I now believe, that Newton's Relative Spaces were of type (oK)^2.

(Examples: Bob Hawke is the Prime-Minister.
Andromeda has the same shape as our galaxy.
The universe has the topology that the last speaker said it had.)

Chris Mortensen and Graham Nerlich in their [1983] yield some ground to the sort of considerations outlined above. The Kantian argument from the existence of certain topological properties of space to the absolute nature of space is unjustifiable, we are told. But evidently 'realism' if not 'absolutism' is justified. 'There is nothing about the case to suggest that the spaces of Galilean Frames (Newton's relative spaces) . . . are not real. They are parts . . . of spacetime; . . . The abolition of absolute space is not inevitably the abolition of space. Nor is it the triumph of relationism.'

If by 'real' we mean 'not abstract', all this is not much of a triumph for realism either, for abstract relative times and spaces can have topological properties also. If by 'real' we mean 'existent' and we allow that abstractions may exist, then any sort of time or space can be real. The existence of a Leibniz space, for example, is sufficient for the truth of 'realism' and the realist thesis is trivialised. Mortensen and Nerlich are correct to insist that 'relative' is not to be conflated with 'non-existent'. But it is important also to distinguish both of these from 'abstract' and 'relational'. 'Time' and 'Space' could yet be non-relational terms referring to abstractions.

To ignore these distinctions is more than a mere logical fallacy. It is to conflate two quite distinct ontological problems as well. One of these concerns the nominalistic programme with respect to abstractions in general and to abstract space, time and space-time in particular. Are some theories which invoke such abstractions *conservative* (to use the terminology of Hartry Field [1980])[5] with respect to nominalised physics? If so, should we regard such theories as false but useful as Field regards number theory? Or should we regard such theories as true and perhaps, because of their conventional origins, *analytically* true? That seems to be a metaphysical problem at least in the sense that it is a problem whose solution is independent of what physical facts are to hand.

The other ontological problem is not of that sort. It is a problem of physics. It concerns whether there is an omnipresent and/or eternal causally efficacious space, time or space-time whose existence must be assumed to explain such causal relations as are known otherwise to obtain.

To sum up: to point out that an item has a contingent property, whether or not the property varies and whether or not it is a topological or geometrical property, is not to entail or even to give evidence in favour of the item being in any particular one of the six types outlined above. Thus in no sense of 'absolute' can the presence of such properties be said to be evidence for the item being absolute. In all six types there are terms which describe or refer to items which are orientable, n-dimensional, Euclidean, Reimanian and so on—and terms from all these different categories can feature in physical explanation.

---

[5] I am indebted to Hartry Field for some discussion on this matter, but any errors in this are my own.

To fail to recognise the possibility of spaces, times and space-times of all six types is to risk the conflation of two quite distinct ontological issues—one which concerns the existence of abstract objects and one which concerns the existence of causally efficacious substrata.

University of Queensland, 1986

REFERENCES

FIELD, HARTRY [1980]: *Science Without Numbers*, O.U.P.

HINCKFUSS, IAN [1975]: *The Existence of Space and Time*, O.U.P.

MACKIE, JOHN [1981]: 'Three Steps Towards Absolutism' in Swinburne, R. (ed.) *Space, Time and Causality*, D. Reidel.

MORTENSEN, CHRIS [1977]: Critical Notice of Hinckfuss [1975] and Nerlich [1976]. *Australasian Journal of Philosophy*, 55, pp. 149–157.

MORTENSEN, CHRIS, and NERLICH, GRAHAM [1983]: 'Spacetime and Handedness', *Ratio*, 25, pp. 1–13.

NERLICH, GRAHAM [1973]: 'Hands, Knees and Absolute Space', *Journal of Philosophy*, 70, pp. 337–51.

NERLICH, GRAHAM [1976]: *The Shape of Space*, Cambride U.P.

NERLICH, GRAHAM [1979]: 'What Can Geometry Explain?', *British Journal for the Philosophy of Science*, 30, pp. 69–83.

RENNIE, MALCOLM [1974]: *Some Uses of Type Theory in the Analysis of Language*, Research School of Social Studies, Australian National University.

SKLAR, LAWRENCE [1974]: 'Incongruous Counterparts, Intrinsic Features and the Substantiviality of Space', *Journal of Philosophy*, 71, pp. 277–290.

Brit. J. Phil. Sci. 39 (1988), 193–210 *Printed in Great Britain*

What is a Law of Nature? A Humean Answer*

PETER URBACH

Introduction
1 *Subjunctive Conditionals and Laws*
2 *Explaining Subjunctive Conditional Reasoning*
3 *Lawlikeness*
4 *Lawlikeness as a Real Feature of the World*
5 *A Subjectivist View of Laws*

INTRODUCTION

True universal generalisations are often categorised either as accidental truths or as laws. Frequently-cited examples of the former are: 'All my friends speak French', 'All the coins in my pocket on VE Day are nickels', and 'All moas die before reaching the age of 50'. The following, on the other hand, are often supposed, if true, to be laws of nature: 'All metals expand on heating', 'In a closed physical system, the total energy is constant', and Newton's 'Laws'.

Views differ about the basis of the intuitive distinction between the two types of generalisation. Some philosophers claim that laws, unlike accidental generalisations, express a 'physical' or 'natural' necessity; others hold that the difference between the two types of generalisation rests in part on some aspect of our personal attitude towards them. The first of these, which I call the objectivist view, has the longer tradition and is probably the more popular; the second is a subjectivist theory, often associated with David Hume.

In this paper, I shall first propose a theory to account for some intuitive differences between laws and accidental generalisations and for some characteristic patterns of reasoning connected with each. I shall then consider whether the distinction between the two sorts of generalisation is best interpreted along objectivist, or subjectivist lines. I shall conclude in favour of the latter, on the grounds that the purported notions of physical necessity or nomic connection which objectivists have hitherto advocated are no less obscure than the intuitive idea of lawlikeness they were supposed to explicate.

* I am most grateful to David Armstrong, to Colin Howson and to Peter Milne for valuable advice and numerous suggestions which led to the improvement of this paper. Of course, I alone am entirely responsible for the end result.

I SUBJUNCTIVE CONDITIONALS AND LAWS

1(a) *Laws and inductive support for subjunctive conditionals.* It is often affirmed that laws support subjunctive conditionals while accidental generalisations do not, and that this is the crucial difference between the two. I think that if slightly amended, this is correct, and shall base my own view on it.

I shall consider generalisations of the form 'All As are Bs' and related subjunctive conditionals like 'If j were A then it would be B'. (I shall abbreviate such statements to $Aj \rightarrow Bj$. The conditional is said to be counterfactual when its antecedent is false.) The thesis we are considering is that only laws support such subjunctive conditionals.

This rule, I suggest, is not strictly correct, for some accidental generalisations also provide such support. For example, suppose coins are selected at random from a hat containing two-headed and two-tailed coins, and that they are then tossed. Consider now the, let us say, true statement: 'All throws with the coins drawn at random from this hat land heads'. If only a few coins were sampled, then probably most people would agree that this statement is accidental: some two-tailed coins could have been drawn, in which case tails would occasionally have been observed, and the statement would have been false. But although accidental, the generalisation, when combined with information about the composition of the hat, implies that only double-headed coins were selected and, hence, supports subjunctive conditionals of the kind: 'If the third coin drawn from the hat were tossed again, it would again land heads'. This example is rather contrived, but some real theories fit the same pattern. For instance, Kepler's theory of elliptical planetary orbits states that all planets with angular displacement e at time t have a radial coordinate $r = f(e)$ at t. This is now thought to be only accidentally true (or, at least, almost true) on the grounds that the positions, sizes and energies of the various bits of matter resulting from the putative primordial explosion are not fully determined by the conditions of that explosion. The planets, if there had been any, might each have had a total energy that would have produced non-elliptical paths. But even so, Kepler's so-called law appears to support subjunctive conditionals like: 'If the planet Mars were at point p at time t, then it would be at point p′ at t′'.

These examples show that laws and accidental generalisations cannot be distinguished purely by whether or not they support subjunctive conditionals. However, those accidental generalisations that support subjunctive conditionals seem to do so because they, so to speak, contain a lawlike element. Indeed, the nomic connection which laws are supposed to describe seems to be more explicitly expressed or encapsulated by the modal language of subjunctive conditionals. My plan, accordingly, will be to investigate the subjunctive conditional and to propose a theory of its truth-conditions; I shall then show the considerable explanatory power of that theory, and advance a view on how laws are able to support subjunctive conditionals. Finally, I shall discuss whether the truth-values of subjunctive conditionals and the lawlikeness of generalisations are determined solely by the state of the external world, or whether, in part, by the state of the mind.

1(b) *Subjunctive conditionals as probability relations.* Subjunctive conditional statements are not simply material implications, for they are not automatically true if their antecedents are false. (E.g., 'If I had started this paper yesterday, I would have finished it today' is false.) Unlike material implications, subjunctive conditionals express some kind of link or necessity which operates between the antecedent and the consequent events, whether or not the former is actualised. In this respect, they resemble conditional probabilities, for the value of $p(Bj/Aj)$ is also independent of whether Aj is true. This and other reasons to be explored herein prompt me to propose that part of the meaning of the subjunctive conditional $Aj \underset{s}{\rightarrow} Bj$ is that $p(Bj/Aj) = 1$.

There must, however, be a further aspect to subjunctive conditionals, for they ought to obey the rules of modus ponens and modus tollens, that is, $Aj \underset{s}{\rightarrow} Bj$ and Aj should imply Bj, and the occurrence of Aj in the absence of Bj should refute $Aj \underset{s}{\rightarrow} Bj$. For this reason, I propose that the subjunctive conditional $Aj \underset{s}{\rightarrow} Bj$ should be understood as asserting not only $p(Bj/Aj) = 1$, but also the material implication, $Aj \rightarrow Bj$.

This interpretation has the advantage that it does not make every contrary-to-fact conditional true, as they would be if understood simply as material implications. For if $p(Bj/Aj) = 1$, then $p(\sim Bj/Aj) \neq 1$; so, on the proposed interpretation, if $Aj \underset{s}{\rightarrow} Bj$ is true, $Aj \underset{s}{\rightarrow} \sim Bj$ is false. Another trivial result is that $P \underset{s}{\rightarrow} P$ is always true, provided $p(P) > 0$. In the following sections, I shall describe how various other features of subjunctive conditional reasoning may be accounted for, and I shall then extend this analysis of the subjunctive conditional to the idea of a law. Finally, I shall consider whether or not lawlikeness should be regarded as a purely objective feature of the world.

2 EXPLAINING SUBJUNCTIVE CONDITIONAL REASONING

2(a) *The principle of contraposition.* There has been some discussion over whether the principle of contraposition applies to subjunctive conditionals. I believe it does, for if j being A in some sense necessitates its being B, so that $Aj \underset{s}{\rightarrow} Bj$ is true, then surely its not being B ensures, by the same necessity, that it would not be, or could not have been A. So, for instance, the following pair of subjunctive conditionals ought to be equivalent:

If this system were closed, then its total energy would be constant.
If the total energy of this system were not constant, then it would not be closed.

The principle of contraposition seems essential to scientific reasoning, and it explains why a well-established theory asserting that $P \underset{\cdot}{\rightarrow} Q$ is often used as the rationale for a diagnostic test of the falsity of P, it being argued that if Q were false, then P would be false too. For example, if a gas contains carbon dioxide, it would turn limewater cloudy; this familiar theory justifies the limewater test for carbon dioxide, on the grounds that if the limewater were to remain clear, then carbon dioxide would be absent. It is surprising, therefore, that David Lewis regards contraposition as a fallacy when applied to subjunctive conditionals. He has offered the following allegedly invalid argument as a counterexample:

If Boris had gone to the party, Olga would still have gone.

∴, if Olga had not gone, Boris would still not have gone.

Lewis argued thus: under the condition that 'Olga would have gone all the more willingly if Boris had been there', the premise is true. However, if 'Boris wanted to go, but stayed away solely in order to avoid Olga', then the conclusion is false. ([1973], p. 35).

But why should Lewis's condition render the premise true? The only reason I can think of is that if Olga were all the more willing to go, she would actually go. Then, assuming the transitivity of subjunctive conditionals, one could infer the premise, namely 'if Boris had gone to the party, then Olga would have gone too'. But the rest of Lewis's argument rules out transitivity for subjunctive conditionals; for he maintains that the conclusion of the above argument is false and that, in fact, if Olga had not gone, Boris would have gone. But we know from the argument's premise that if Boris had gone, Olga would have gone, whence, by transitivity, if Olga had not gone, then she would have gone, which clearly is false.

So Lewis must either give up his claim that the conclusion of the argument is false under the conditions he imagines, in which case the contraposition principle is unscathed. Or he must deny transitivity (which he does, as we shall see), in which case, he would be unable to claim that the premise of the argument is true, and the contraposition principle would again be unimpaired.

Since I think it plausible for contraposition to apply to subjunctive conditionals and since, if I have understood it right, Lewis's argument fails, it seems to me an advantage of my proposal that it preserves and explains the principle. The explanation proceeds as follows. Suppose $S \underset{\cdot}{\rightarrow} Q$. Then, according to the suggested interpretation, this implies that $p(Q/S)=1$. But,

$$\begin{aligned} p(Q/S) &= 1 - p({\sim}Q/S) \\ &= 1 - p(S/{\sim}Q)p({\sim}Q)/p(S) \\ &= 1 - [1 - p({\sim}S/{\sim}Q)]p({\sim}Q)/p(S). \end{aligned}$$

And if $p(Q/S)=1$, then, provided that neither $p(S)$ nor $p({\sim}Q)$ are zero, $p({\sim}S/$

$\sim Q) = 1$. Clearly, the material implication part of the subjunctive conditional obeys the contraposition rule. Hence, $\sim Q \underset{\cdot}{\rightarrow} \sim S$, as is required.

2(b) *Strengthening the antecedent of a subjunctive conditional.* Another property one might anticipate is that a true subjunctive conditional remains true when its antecedent is strengthened. For instance, if this metal would expand if it were heated, then it would also expand if it were heated in a magnetic field, say. In general, if $Aj \underset{\cdot}{\rightarrow} Bj$ is true, so is $Aj \;\&\; C \underset{\cdot}{\rightarrow} Bj$.

Lewis (p. 31), however, regards this as fallacious, a claim he attempted to demonstrate by means of what to me are unconvincing counterexamples (p. 10). For instance:

If the USA threw is weapons into the sea tomorrow, there would be war.

$\therefore$, if the USA and the other nuclear powers all threw their weapons into the sea tomorrow there would be war.

Lewis contended that the premise of this argument is true but its conclusion false, and so concluded that the inference is invalid. But surely, the premise is only true if it is understood as predicting war under the condition that the USA, but not all other powers discarded their weapons. If the premise is more fully spelled out, so that it includes that tacit condition, it becomes clear that the conclusion is not deduced by strengthening the antecedent. So Lewis's illustration is no counterexample to the rule of inference we are discussing. Similar objections apply to Lewis's other putative counterexamples (e.g., 'If Otto had come it would have been a lively party; but if both Otto and Anna had come it would have been a dreary party').

I see it, therefore, as an advantage of the present theory of subjunctive conditionals that it requires the principle of strengthening the antecedent. That it does so can be proved, for, first, if $p(Bj/Aj) = 1$, then, provided $p(Aj \;\&\; C) > 0$, $p(Bj/Aj \;\&\; C) = 1$; and, secondly, the material implication part of a subjunctive conditional clearly obeys the principle. The proof of the first part (for which I am indebted to Peter Milne) is as follows. I shall simplify the working by writing AB for A & B, A for Aj, and $\bar{A}$ for $\sim A$.

Suppose $p(B/A) = 1$, then $P(AB) = p(A)$.
Hence, $p(ABC) + p(AB\bar{C}) = p(AC) + p(A\bar{C})$.
But since $p(A\bar{C}) = p(AB\bar{C}) + p(A\bar{B}\bar{C})$,
we have $p(ABC) + p(AB\bar{C}) = p(AC) + p(AB\bar{C}) + p(A\bar{B}\bar{C})$.
Hence, $p(ABC) = p(AC) + p(A\bar{B}\bar{C})$.
Therefore, $p(B/AC) = 1 + p(A\bar{B}\bar{C})/p(AC)$.

We can show that $p(A\bar{B}\bar{C}) = 0$, for since $p(AB) = p(A)$, $p(A\bar{B}) = 0$; hence $p(A\bar{B}\bar{C}) = 0$. Thus, provided $p(AC)$ is non-zero, $p(B/AC) = 1$, which is what we set out to prove.

2(c) *The principle of transitivity.* Subjunctive conditionals also seem to obey a rule of transitivity, in that if $C \underset{s}{\rightarrow} A$ and $A \underset{s}{\rightarrow} B$, then $C \underset{s}{\rightarrow} B$ may be inferred. So, for example, 'if this metallic object were heated, it would expand' and 'if it expanded, it would shatter its container' together imply that if the object were heated it would shatter its container. This principle seems to underlie much thinking in science, particularly in regard to its technological applications. A safety device, for instance, is designed on the calculation that if some fault occurred, this would trigger a certain response, which in turn would set off another, and that the end result of such a sequence would be for the fault to be rectified or mitigated. In view of the various intervening mechanisms, the scientist would then be prepared to argue that if a fault were to arise, then it would be rectified or mitigated.

Lewis has, however, argued that the principle is fallacious, on the basis of alleged counterexamples such as:

If Otto had gone to the party, then Anna would have gone.
If Anna had gone, then Waldo would have gone.

$\therefore$, if Otto had gone, then Waldo would have gone.

Lewis (p. 32) explained that 'Otto is Waldo's successful rival for Anna's affections. Waldo still tags around after Anna, but never runs the risk of meeting Otto. Otto was locked up at the time of the party, so that his going to it is a far-fetched supposition; but Anna almost did go'. Under these conditions, Lewis stated, the premises are true and the conclusion false.

Presumably Lewis would argue that since Waldo never runs the risk of meeting his rival, if Otto were at the party, Waldo would not be, which we may summarise as $O \underset{s}{\rightarrow} {\sim} W$; hence the conclusion of the argument is false. The truth of the first premise is secured by Otto's success in winning Anna's affections, so that $O \underset{s}{\rightarrow} A$. The second premise is presumably held to be true on the grounds that Waldo pursues Anna, except when Otto is around, i.e., $A \& {\sim} O \underset{s}{\rightarrow} W$. And since Otto is tied up on the day of the party, we are presumably meant to infer the second premise, $A \underset{s}{\rightarrow} W$. This inference seems to me invalid. It is presumably not meant to be in the form $A \& {\sim} O \underset{s}{\rightarrow} W$, ${\sim} O$, therefore $A \underset{s}{\rightarrow} W$, which would not generally be accepted as valid, and is certainly not valid according to Lewis's account of subjunctive conditionals. The inference might perhaps be valid if O is, in some sense, physically impossible, so that ${\sim} O$ is physically necessary. But in that case, the status of subjunctive conditionals in which O is part of the antecedent is problematic. 'If water were composed of nickel and zinc, then . . .' or 'If I were my own grandfather, then . . .', or (when taken literally) 'If I were you, then . . .' exemplify a type of statement of doubtful meaning which do not seem to me to have a definite truth value, and this is reflected in the theory advanced in this paper by the fact that $A \underset{s}{\rightarrow} B$ is undefined when $p(A) = 0$.

Therefore, I see no reason to reject the plausible idea that subjunctive conditionals are transitive, and regard it an advantage of the present theory that it incorporates such a rule. To show that it does, we must prove that if $p(A/C)=1$ and $p(B/A)=1$, then $p(B/C)=1$. (Clearly, transitivity holds for the corresponding material implication.) The proof (which I owe to Colin Howson) proceeds thus:

Suppose $p(A/C)=1$ and $p(B/A)=1$.
Then $p(AC)=p(C)$ and $p(BA)=p(A)$.
Since $p(AC)+p(\bar{A}C)=p(C)=p(AC)$,
it follows that $p(\bar{A}C)=0$.
Similarly $p(\bar{B}A)=0$.
Now, $p(\bar{B}C)=p(\bar{B}CA)+p(\bar{B}C\bar{A})$
$= \quad 0 \quad + \quad 0$
hence, $p(\bar{B}C)=0$.
And since, $p(C)=p(BC)+p(\bar{B}C)$ and $p(C)=p(BC)$,
it follows that $p(B/C)=1$, provided $p(C)>0$.

Hence, subjunctive conditionals are transitive. These conditionals also obey a number of other plausible and scientifically useful principles of inference. For instance, from $A\rightarrow_s B$ and $C\rightarrow_s D$, $A\&C\rightarrow_s B\&D$ follows; and if $A\rightarrow_s B$ and $A\rightarrow_s C$, we may infer $A\rightarrow_s B\&C$ (the usual provisos about the antecedents not having zero probability holding).

2(d) *Simplifying Disjunctive Antecedents.* Suppose that $AvB\rightarrow_s C$ implies $A\rightarrow_s C$ (call this the 'Simplification Principle'). Presumably, $A\rightarrow_s C$ is equivalent to $((A\&B)v(A\&{\sim}B))\rightarrow_s C$ which, according to this principle, entails $A\&B\rightarrow_s C$; hence, $A\rightarrow_s C$ entails $A\&B\rightarrow_s C$. (Ellis [1979], p. 68). Thus if the simplification principle is valid, so is that of the strengthening of the antecedent, with which we dealt in 2(b). Since Lewis denies the latter, he must deny the former also and, indeed, he does. Brian Ellis [1979] also denies the simplification principle, though for slightly different reasons. Nevertheless, both Ellis and Lewis believe to be valid, the informal version of the principle, i.e., that 'If A or B occurred, then C would happen' implies 'If A occurred, then C would happen'. Hence, they must deny that subjunctive conditionals with disjunctive antecedents are of the form $AvB\rightarrow_s C$. Both tentatively suggest that $A\rightarrow_s C$ & $B\rightarrow_s C$ would better express such subjunctive conditionals, though as Ellis conceded, there is little more than the need to accommodate the simplification principle itself to justify such a re-interpretation.

The theory of subjunctive conditionals advanced here does not require any such ad hoc re-interpretation. For on that theory, $A\rightarrow_s C$ is a logical consequence of $AvB\rightarrow_s C$. To prove this one needs to show that $p(C/A)=1$ is entailed by $p(C/AvB)=1$, and that $A\rightarrow C$ follows from $AvB\rightarrow C$. The latter is obvious. The proof of the former is as follows:

If $p(C/AvB) = 1$, then (provided $p(AvB) > 0$) $\frac{p(C(AvB))}{p(AvB)} = 1$

$$\therefore \quad p((CA) \text{ v } (CB)) = p(AvB)$$
$$\therefore \quad p(CA) + p(CB) - p(ABC) = p(A) + p(B) - p(AB)$$
$$\therefore \quad p(AB) - p(ABC) = p(A\bar{C}) + p(B\bar{C})$$
$$\therefore \quad p(AB\bar{C}) = p(A\bar{C}) + p(B\bar{C})$$
$$\therefore \quad p(A\bar{C}) + p(B\bar{C}\bar{A}) = 0$$
$$\therefore \quad p(A\bar{C}) = 0$$
$$\therefore \quad \text{if } p(A) > 0,\ p(C/A) = 1.$$

Thus the simplification principle necessarily applies to subjunctive conditionals and there is no need to interpret the disjunctive antecedents of such conditionals in any other way than the natural one.

2(e) *A possible objection rebutted.* The interpretation I am offering for subjunctive conditionals evidently accounts for some of their characteristics. However, before proceeding I wish to consider a plausible objection, put to me by Peter Milne, which argues that while $Aj \underset{\mathrm{I}}{\rightarrow} Bj$ rules out the possibility of j not being B when it is A, the probability statement '$p(Bj/Aj) = 1$' does not; hence, the subjunctive conditional cannot be glossed in the manner I suggested.

The following example shows why one might sometimes regard a specific event with unit probability as not being nomically necessary. A lottery has an uncountably infinite number of tickets, one of which is drawn at random. The probability of drawing any particular ticket must be zero; hence its probability of not being chosen is necessarily 1. But despite this maximum probability of not picking any given ticket, we should not say that the ticket could not possibly have been drawn, since the structure of the experiment ensures that it could have been.

This experiment is, of course, fanciful; there are no such lotteries. The question is whether analogous, authentic examples exist. Ordinary scientific reasoning might suggest that they do. For instance, inductive logicians investigating a parameter such as a distance, a velocity, or a temperature, frequently proceed by postulating a prior density-distribution over a continuous range of values of the parameter, and then calculate a corresponding posterior distribution in the light of experimental evidence. Although each parameter value in the range might have zero probability, both before and after the experiment, the manner of proceeding suggests that each is being regarded as a physically possible candidate for the truth. Similarly, scientists frequently assume physical quantities to be related by a continuous function, thereby suggesting, apparently, that each real number in a certain interval is a genuine possibility for the values of those quantities.

It seems to me, however, that this is not the right conclusion to draw from the methods described. The reason why densities are often defined for physical parameters over continuous ranges, I submit, lies in the mathematical convenience and conceptual simplicity which such a device brings and does not indicate a commitment to the physical possibility of each element in the continuum. For suppose the parameter in question is a length, and consider the possibility that the true value is *exactly* t meters. This means that the length we are interested in is a multiple t of the length of the standard meter preserved in Paris. But this standard meter, according to the molecular theory, does not have a fixed distance from end to end, so without some further distinction, one could not hold that a length is *precisely* t times as long as the standard meter. But it is rather difficult to see how the infinite precision implied by the theory could be made meaningful. Even if one could fix some moment in the life-time of the Paris meter when it acts as a standard, there is the question of which part (if we may speak in such terms) of the terminal molecules count as its end points, and the problem of defining an infinitely precise temporal moment. These difficulties suggest that to regard every real number in a continuous range as a genuine possibility for an unknown length is to imagine a theoretical precision which is simply not available and is not really assumed by scientists. I believe the same can be said about any physical parameters such as times, velocities, temperatures, etc.

This argument, I contend, shows that the scientific practice of handling continuous functions of physical variables does not imply that each value which it 'could' have is being entertained as a genuine possibility. The practice should be regarded rather as an idealisation which greatly simplifies calculation and conceptualisation. Hence, in reply to the above objection, I claim that $p(B_j/A_j)=1$ does not really admit the possibility of j not being B when it is A.

3 LAWLIKENESS

Lawlike statements share all the characteristics I have considered in connection with subjunctive conditionals. Like subjunctive conditionals, they may be, but are not necessarily, true even when their antecedents are not instantiated, and they appear to obey the same rules of inference, such as the rules of modus ponens, modus tollens, transitivity and contraposition. This suggests that laws are simply generalised subjunctive conditionals; in other words, 'All As are Bs' is a law just in case, for every x, $p(Bx/Ax)=1$ and $\forall x(Ax \rightarrow Bx)$. (Of course, the variable x must be understood as ranging over actual objects in the real world, not possible objects in possible words.) On the other hand, while an accidental generalisation implies the universal material conditional, it does not imply the corresponding conditional probability. Laws thus understood support subjunctive conditionals by logically implying them.

4 LAWLIKENESS AS A REAL FEATURE OF THE WORLD

I have yet to consider what meaning should be assigned to the probability by which the subjunctive conditional has been interpreted. Whether it has a purely objective, or a partly subjective denotation will determine whether the interpretation of lawlikeness I have offered should be placed into the objectivist or the subjectivist camp. It has been widely argued that the former would be the more satisfactory, and a number of attempts have been made to describe the feature of the world corresponding to the necessary connection that laws, but not accidental generalisations, supposedly express. I shall review the principal theories of this sort; my conclusion will be that they all fail in their purpose and that, at present, there is no alternative to a Humean interpretation of natural laws.

4(a) *Relations between universals.* Dretske [1977] proposed an objectivist theory of lawlikeness, according to which laws are not merely generalisations of matters of fact but are relations between universals, or properties. In this respect, the theory I am offering agrees with it. There are, however, in my view, two significant differences. First, Dretske's theory has much less explanatory power. Secondly, the postulated nomic relation is as obscure as the intuitive idea it is supposed to illuminate, while the type of lawlike connection advanced in this paper has a relatively clear interpretation (see section 5). Dretske represented a nomological statement connecting objects which are F with those with the character G in the form: F-ness→G-ness. The arrow used here is not the symbol for material implication; according to Dretske, it expresses 'a link or connection between the respective qualities', which could also be rendered as 'yields', as for instance if 'all men are mortal' is a law, we might say that 'humanity yields mortality' (p. 253). Attempting to elucidate this further, Dretske said that 'F-ness is linked to G-ness; the one property yields or generates the other in much the way a change in the thermal conductivity of a metal yields a change in its electrical conductivity' (p. 264). This link, or connection, or process of yielding and generating enables one to 'understand the modal character of laws', namely a law's capacity to tell us 'what (in some sense) must happen, not merely what has and will happen'. (p. 263)

Links and connections are unmysterious when they are material ties between physical objects, and abstract links, in the form of functional relationships between physical objects or events, are also easily comprehended. However, the linkage Dretske envisaged connects properties and is more difficult to understand. He attempted to make the idea more accessible by explaining that when two properties are nomically related, the one 'yields or generates' the other. But this seems to me to add nothing to the simple assertion that events of the one kind are invariably succeeded by those of the other. And to explain this yielding as being much the same as occurs when a change in the thermal conductivity of a metal yields a change in its electrical conductivity is to resort to a mystery to clear up an enigma. After all, it is precisely the nature of the connection between the entities referred to by laws such as the one concerning electrical and thermal conductivity that his example was intended to explicate.

4(b) *Lawlikeness explained through possible worlds.* Dretske also attempted to shed light on the nature of the special modal character of laws by invoking the idea of possible worlds. Thus he claimed that because laws are relationships between properties, they '*go beyond* the sets of things in *this* world... Their scope extends to those possible worlds in which the extensions of our terms differ but the connections between properties remains invariant'. Hence, laws support counterfactual conditionals. Accidental generalisations do not, for they tell us 'absolutely nothing about those possible worlds in which there are *additional* Fs or *different* Fs' (p. 266). This distinction sounds informative but it is not. For suppose we had access to the various possible worlds and could decide for each of them whether or not a generalisation was true. We would still not know whether the generalisation was a law. To find that out, we would also have to show that the generalisation held in all those worlds sharing the same 'connections between properties'. Since the only connections between properties Dretske referred to are nomic ones, we are left with the unhelpful claim that a law is a statement that is true in all worlds in which it is a law.

Popper too has been round this circle. He suggested that a universal statement is a law if it holds in all worlds which differ, if at all, from ours with respect only to initial conditions. ([1959], p. 433). But what is it for a world to differ from ours *only* with respect to initial conditions? Clearly it means that the world departs from this one in its initial conditions but that it is governed by the same laws. So, as has been pointed out by, for example, Nerlich and Suchting ([1967]), Popper's definition is circular in a way which makes it useless for characterising laws; the definition of a law uses the notion of worlds differing with respect to initial conditions only, and this has to be explicated via the notion of a law.

This circularity is not surprising. Possible worlds are states of affairs whose existence is not barred by logic. A law cannot be true in all those worlds, unless it is a law of logic. A physical law must be false in some possible worlds. But which? Since we have no relevant information about such worlds, except that the generalisation in question is false within them, we seem bound to arrive at the circular definition that a law is true in all those worlds where it is true. This point is made lucidly by Cohen in his [1980].

4(c) *Armstrong's and Tooley's theories.* A similar theory to Dretske's has been proposed by Tooley ([1977]) and, independently and in greater detail, by Armstrong ([1978] and [1983]). Thus Armstrong (p. 172) argued that 'laws of nature are diadic relations of necessitation (or probabilification) holding between universals'. Tooley similarly maintained that 'All Ps are Qs' is a natural law when the ordered pair <P,Q> are nomically related. Such relations subsist 'among universals and nothing else', that is, they cannot be 'analysed in terms of properties of, and relations among, particulars' (Tooley, p. 679). However, Tooley (p. 673) added that the proposition that P and Q are nomically related implies the generalisation that everything P is also Q. This is evidently an extra condition on the idea of nomic relatedness.

As Tooley himself conceded (p. 683), it is easy to gain the impression that his theory makes no real advance in the analysis of lawlikeness, for it leaves the nature of the proposed nomological relationship unexplained; the relationship turns out to be merely that linkage in virtue of which the corresponding generalisation expresses a law. This spreads as much light on the question as the infamous dormitive virtue casts on the soporific effect of opium. Indeed, evidence that 'nomological relationship' is just a new name for an old unknown comes from Tooley's treatment of laws of the form 'All Ps are not Qs'. Persuaded that negative properties are 'suspect' and so cannot enter into nomological relations, he concluded that 'another relation has to be introduced to handle laws of this form: the relation of *nomic exclusion*' (Tooley, p. 676). But nothing further is vouchsafed about this relation, except that it is the one holding when 'All Ps are not Qs' is a law.

One advantage which Tooley saw in his analysis is that it locates the items bearing a nomological relation one to another in universals, or properties. And this, according to Tooley, means that his explication of the lawlikeness of a statement says more than that it is lawlike. He claimed, too, that one is compelled to regard lawlikeness as a characteristic of properties, because an analysis based on particulars cannot comprehend laws with no positive instances, such as laws governing the reactions of substances which are never prepared. Tooley regarded an analysis of laws in terms of possible worlds or subjunctive conditionals as almost bound to be circular (p. 670) and he concluded that it must be 'facts about *universals* that serve as the truth-makers for basic laws' (p. 672). But what facts? What is it about some universals that constitutes their nomic relatedness which others lack? How can universals be related, other than by the normal kinds of relations holding between their instances? These questions seem not to be answered by the advocates of this view, and I can only conclude that far from lifting the mystery from lawlikeness, they merely proliferate new difficulties, more intractable than the first.

A second claim of Tooley's, however, suggests that there is more to nomological relationships than the sparse picture painted so far. He argued that 'one of the merits of the present view is that it does make possible an answer to the epistemological question' of 'how one determines whether a given nomological relation holds among specific universals' (p. 688). Tooley's confirmation theory is a Bayesian one, interest being focused on the posterior probability of a theory. For this to have a positive value, the theory's prior probability must exceed zero, and it was Tooley's claim that this 'seems reasonable if relations among universals are the truth-makers for laws, but not if facts about particulars are the truth-makers' (p. 693). On Tooley's account, a law is 'in a sense, an "atomic" fact, and it would seem perfectly justified, given standard principles of confirmation theory, to assign some non-zero probability to this fact's obtaining'. On the other hand: 'If . . . it is facts about the particulars falling under a generalisation that make it a law, then, if one is dealing with an infinite universe, it is hard to see how one can be justified in assigning any non-zero probability to a generalisation, given evidence concerning only a finite number of instances' (p. 693). There are, in my view, two reasons for rejecting this argument. First, accidental generalisations exist to which one is clearly justified in assigning a non-zero probability. Example: 'All trials consisting of 1 million independent, consecutive throws with this symmetrical and homogeneous coin contain at least 2 per cent heads'. This generalisation is presumably accidental but it is, nevertheless, credible, even when none of its instances has been verified. This is particularly clear when there is evidence suggesting that the coin is unlikely to survive more than, say, one hundred million throws. Secondly, Tooley's argument is not consistent with his characterisation of laws, for this, quite reasonably, stipulates that the proposition that P and Q are nomically related logically entails that all things which are in fact P are also Q. But then the former does *not* have the higher probability.

A possible defence of Tooley, which David Armstrong suggested to me, and which might reflect Tooley's position more closely, is to say that the statement whose probability is supposed to be zero is not merely a generalisation, G, but consists of the further assertion that G expresses no law. The accidental generalisation could then be written as the conjunction $G \& \sim L$. But if this has a zero probability, then either $p(G)=0$ or $p(\sim L/G)=0$. As we have seen, Tooley in effect rules out the first, since laws with non-zero probabilities logically imply the corresponding generalisations. The second seems unacceptable, for it would make it absolutely certain of any generalisation that it is lawlike. I am obliged to conclude then that Tooley's attempt to give substance and clarity to the notion of relations between universals is not successful.

In Armstrong's view, however, the qualms I have aired about the comprehensibility of the notion of nomic relations between universals is misplaced. He has argued that the idea of such relations is simply a generalised version of the causal necessity linking particulars, which we already apprehend with great clarity by direct intuition. In this respect, my own view coincides with Armstrong's. We differ, however, in how we see the basic connection between particulars. According to Armstrong, each of us possesses 'a primitive notion of necessitation' which we can apply without difficulty in determining particular causal relations. For instance, when a small bit of stuff is dropped into a glass of liquid and it immediately explodes, we have no difficulty in recognising that the explosion was produced or was necessitated by the earlier action. However, the nature of this necessitation is inexplicable.

> The inexplicability of necessitation just has to be accepted. Necessitation, the way one Form (universal) brings another along with it as Plato put it in the *Phaedo* . . . is a primitive, or near primitive, which we are forced to postulate. . . . (It was perhaps a shortcoming of the original Dretske-Tooley-Armstrong formulation that this was not made evident.) We must admit it in the spirit of natural piety, to adopt Samuel Alexander's phrase. (Armstrong, [1983], p. 92)

No doubt we often do apprehend causal connections without difficulty. However, our easy commerce with the notion when we are in an unreflective mood does not mean that our reasoning obeys no discoverable rules or that our idea of causation is undeserving or incapable of further elucidation. On the first point, there are indications that one's judgement of a causal link is not immediate, for we all recognise that even in cases where an effect is contiguous in time and space to a putative cause, it may in fact have been caused in some other way. Presumably there are rational processes which lead to the exclusion of some possible causes and the admission of others and these are accessible to further investigation. Secondly, our intuitions are not informative on one of the main points at issue, namely, whether the causal necessitation one attributes to certain processes is paralleled by a physical link between external objects, or whether it is merely a subjective, epistemic phenomenon. Indeed, as soon as the question is raised, it becomes apparent that one needs to consider what kinds of external linkages could possibly act as the physical counterparts of our subjective experience, and that any possibilities which are mooted have to be expressed in commonly accepted terms before they can be understood and communicated. This is how intuitively-based notions, such as colour and warmth, have come to be better understood. If a theory of heat were founded upon a physical caloric virtue whose nature were left unexplored, it would, I think, be coolly received. It seems reasonable that inquiries into our sensations of causal connection should observe the same proprieties of precision and clarity demanded of comparable explanatory theories.

4(d) *Probabilities and accidental generalisations.* Some of the examples of accidental generalisations I have cited arise from what are usually regarded as random processes with independent outcomes possessing definite probabilities. Theories such as 'All the throws with this symmetrical and homogeneous coin yield heads' or 'All atomic decays in this sample of radium occur at equally-spaced intervals' seem particularly clear—albeit, rather incredible—generalisations which, if true, are true by accident. It is perhaps not surprising, then, that philosophers (e.g., Giere) interested in the ideas of objective lawlikeness and accidentalness have attempted to substantiate these notions in terms of physical probabilities, where laws are associated with a maximum probability or propensity and accidental generalisations with some lesser value. The principal objectivist account of physical probability is that a probability is a property of an infinite sequence of trial outcomes. If the possible outcomes of a given experiment are $O_1, O_2, \ldots, O_r$, and $n(O_i)$ represents the number of outcomes of the kind O_i appearing in a sequence of m trials, then the probability of O_i in trials of this kind is given, in the frequency theory, by the limit, as m tends to infinity, of the ratio $n(O_i)/m$.

The definition implies that only types of event, not particular events, have a probability, though Salmon and others have tried to harness an objective probability for specific outcomes of a trial by extending the definition. But whatever judgement one forms about these efforts, approaches based on frequency limits cannot help with the issue at hand, even though they are useful in other contexts, for one cannot comprehend a limit that would be reached if *per impossibile* a trial were repeated to infinity, without first understanding counterfactual conditionals and lawlikeness. (The same point applies to the propensity view of probability.)

5 A SUBJECTIVIST VIEW OF LAWS

One cannot, of course, demonstrate conclusively that every attempt to characterise laws purely objectively will fail, but the limitations of the theories reviewed here support this conclusion and, I suggest, point to a failing in the whole approach which views lawlikeness as a real feature of the world. For this reason, it seems to me that the rival Humean, subjectivist, tradition deserves a more sympathetic treatment from philosophers. The theory I have advanced lends itself to a subjectivist interpretation if the probability term, p(Bj/Aj), which was used to express the subjunctive conditional, Aj$\underset{s}{\rightarrow}$Bj, be regarded as a degree of belief or intensity of conviction.

Some controversy surrounds the idea that degrees of belief are probabilities. My opinion is that it at least approximates the truth, and that the various arguments designed to show that the degrees of belief of reasonable people conform to the probability axioms are good ones. However, this is not the place to review these arguments. (They are reviewed, for instance, in Howson and Urbach, [1989].) In any case, there is no necessity to do so, for my proposal does not depend in an essential way on intensities of belief obeying the axioms of probability exactly; they may perhaps be subject to slightly different rules. I am appealing to the theory of subjective probability because it has some plausibility, and, more importantly, because it is a clear and well-known hypothesis.

Our beliefs evidently do appear with varying intensities, going from a complete certainty that a proposition is true, or that an event will occur, to a maximum conviction that it is false, or that the event will not occur.

Moreover, these different intensities of belief in different propositions are universally and immediately experienced. The varying strengths of conditional beliefs, I suggest, are also directly experienced in the same way, when one imagines the conditions being satisfied. Since they are familiar properties, a theory of lawlikeness raised on the notion of degrees of belief at least avoids the degree of obscurity or circularity with which objectivist accounts are afflicted.

This last claim would be torpedoed if one went along with what I believe is an unfortunate tendency to define subjective probabilities in terms of betting behaviour, whereby a person's degree of belief is identified with the odds that he or she would just be prepared to accept in a wager over the truth of a proposition. Such a definition would automatically disqualify subjective probability from clarifying subjunctive conditionals, since it is itself defined in such terms. Most advocates of subjective probability, however, regard decisions made under conditions of uncertainty as reflecting, amongst other things, an actual state of mind or intensity of belief; the significance of betting behaviour is then that, under certain special conditions, it provides a convenient method for measuring these intensities. This, at least, was de Finetti's view: 'the degree of probability attributed by an individual to a given event is *revealed* by the conditions under which he would be disposed to bet on that event'. ([1936], p. 101; my italics). And de Finetti's exposition of probabilities in terms of coherent betting behaviour was put forward 'above all for its clarity' (p. 102), rather than to claim that probabilities are nothing but dispositions to accept wagers.

The prospect of an objective conception of laws cannot be abandoned without regret; if it is partly subjective, the idea loses some of the metaphysical and epistemological significance which it promised. Nevertheless, if one is interested in the structure of scientific reasoning, it is worthwhile discovering the basis on which philosophers and scientists have found it useful to distinguish between the two types of generalisation, and explaining the origins of the various intuitions linked with that distinction. I believe the above account goes some way towards this.

The London School of Economics

REFERENCES

ARMSTRONG, D. M. [1978]: *Universals and Scientific Realism*. Cambridge University Press.

ARMSTRONG, D. M. [1983]: *What is a Law of Nature?* Cambridge University Press.

COHEN, L. J. [1980]: 'The Problem of Natural Laws' in D. H. Mellor (ed.) *Prospects for Pragmatism*. Cambridge University Press.

DRETSKE, F. I. [1977]: 'Laws of Nature', *Philosophy of Science*, 44, pp. 248–68.

ELLIS, B. [1979]: *Rational Belief Systems*. Basil Blackwell.

FINETTI, B. DE [1937]: 'La prévision; ses lois logiques, ses sources subjectives', *Annales de l'Institute Henri Poincaré*, **7**, pp. 1–68. Reprinted in English translation in H. E. Kyburg and H. E. Smokler [1964] (eds.), *Studies in Subjective Probability*. John Wiley & Sons.

GIERE, R. N. [1979]: 'Propensity and Necessity', *Synthese*, **40**, pp. 439–51.

HOWSON, C. and URBACH, P. [1989]: *Scientific Reasoning: the Bayesian Approach*. Open Court Publishing Company.

LEWIS, D. [1973]: *Counterfactuals*. Basil Blackwell.

NERLICH, G. C. and SUCHTING, W. A. [1967]: 'Popper on Law and Natural Necessity', *The British Journal for the Philosophy of Science*, **18**, pp. 233–5.

POPPER, K. R. [1959]: *The Logic of Scientific Discovery*. Hutchinson.

TOOLEY, M. [1977]: 'The Nature of Laws', *The Canadian Journal of Philosophy*, **7**, pp. 667–98.

SALMON, W. C. [1984]: *Scientific Explanation and the Causal Structure of the World*, Princeton University Press.

Brit. J. Phil. Sci. 39 (1988), 211–223 *Printed in Great Britain*

Empirical Tests Are Only Auxiliary Devices*

TYRONE LAI

We test a scientific theory empirically by deducing from it a 'prediction' and comparing it with observation. In Popper's view of science, theories can be refuted this way, that is through tests, but can never be rendered more probable. In his view, even when a theory has passed all the tests to which we have subjected it, it still remains as a conjecture with zero probability (Popper [1962], p192). Popper's view of science includes a form of skepticism. According to Popper, we can propose theories (in answer to problems) and try to refute them; in doing so we *hope* we come closer to the truth, but we will never know whether we have actually done so (Popper [1959], p415; Hattiangadi [1983]).

Suppose we accept Popper's argument that theories cannot be rendered more probable by empirical tests. Do we have to accept his form of skepticism?

I think not. Popper's argument would lead to skepticism only if empirical testing constitutes the *central* mechanism in the evaluation of theories.[1] If it is not the central mechanism; if there is some other way of evaluating theories more important than tests; skepticism may not follow.

In this paper, I argue that tests are but auxiliary devices. Though definitely useful, they are dispensable. There is a more important method of evaluating hypotheses. This method of evaluation I call the method of progressive evaluation, or MPE. MPE is the central method of evaluating hypotheses in science—and in other kinds of empirical investigations. When we apply this method in evaluating an hypothesis, if the hypothesis is true (without ourselves knowing), we will become progressively more certain that it is. If the hypothesis is a good approximation, the investigative strategy, IVS, associated with this method will enable us to improve upon the approximation. If the hypothesis is false, and seriously defective to boot, MPE will allow us to eliminate it. In a word, MPE together with IVS, allows us to uncover knowledge *gradually*. Whether this will satisfy all skeptics is open to question; however, if we do have this method, we will have some indication as to whether we are uncovering knowledge. I will explain MPE and IVS later on in this paper.

---

[1] Theories can be evaluated in other ways. For example, logically, to see whether they are internally consistent.

* I wish to thank an anonymous referee for helpful comments on an earlier version of this paper.

To help make my point that tests are but auxiliary devices, and to explain MPE and IVS, I shall make use of a simple exercise in cryptanalysis (codebreaking: for an introduction to the subject, see Kahn [1967]). Such an exercise provides us with an example of an empirical investigation *in compact form*. It is important that we should have a compact example. Most investigations, including those in science, tend to be untidy, and drawn out over time. In these cases, it is much more difficult to see the actual method used for evaluating hypotheses.

Cryptanalysis is nowadays a sophisticated science/art. Cryptanalysts in their work often have to use specialized techniques. In this paper, we are only interested in the general strategy and the method of evaluating results which they share with investigators in other areas.

Some may question that cryptanalysis can be representative of empirical investigations, particularly science. In cryptanalysis, they will say, the secrets we aim at are already known—to some people. But in science, the secrets we aim at are not yet known to anybody. This objection is quite understandable, but not valid. For, cryptanalysts work not only on ciphers currently in use, they also decipher dead languages. But even more important than this is the very nature of the activity itself. Cryptanalysis is called for only when those who know will not, or are not able to, tell. Thus even when the secrets cryptanalysts are trying to uncover are still known to some people, this fact will not make any difference to their work; they still have to find out these secrets on their own, and to determine on their own to what extent they have succeeded; they cannot expect help from those who know.[2]

Cryptanalysts work very much the same way scientists work. They both have to depend on trial and error, guided sometimes by background information. They both have to evaluate results without direct access to the realities they are interested in.

I now present our example in cryptanalysis, commenting on it as I proceed. In these comments, I shall be drawing attention to the way we evaluate hypotheses, as well as further parallels between science and cryptanalysis. The explanations for IVS and MPE will be given in the comments to Step 4. The central thesis of this paper—that empirical tests are but auxiliary devices—will be explained in the comments to Step 5.

---

[2] Kuhn [1962] in his comparison of normal science to puzzle-solving, seems to imply that secrets that are known to *some* people somehow make those investigations aimed at uncovering them *by those who do not know* less of an investigation, as though these latter could have known, if they had wanted to, even without the investigation. It is easy to arrive at this misconception through games such as jigsaw puzzles and crossword puzzles. In these games, when asked, those who make up the puzzles are able to let the puzzle-solver know what the answers should be. In this sense the puzzle-solver could have known even without the investigation. He could have asked! However, in cryptanalysis, we cannot expect cooperation from those who make up the ciphers. Even in games in cryptanalysis, to imitate real life, we try not to rely on such cooperation. But then, it will be asked, how do we know we have the right results? And this of course is the interesting point, one which ordinary people, and also philosophers, often overlook. *We can direct investigations to uncover secrets, and evaluate, objectively, our results, without having at any point, before or after, direct access to those secrets.*

EXAMPLE IN CRYPTANALYSIS

Cryptogramme:

SBR SBCTU DBCKERVS FCGG WTTCXR SFH FRRJD YTHE SHUWI

Comments:

In the present instance, the cryptogramme is given. In other instances, it may first have to be discovered.

In science, the 'cryptogramme' always has to be discovered. Moreover, it may not occur in one tidy bundle. Does the variation of the brightness of Venus belong to the same cryptogramme as its motion? Copernicans thought yes; Aristotelians thought no.

Analysis:

Step 1. Assumption A1: Plaintext is in English.
A2: One ciphertext letter stands for one and the same plaintext letter (simple substitution cipher).
Suspicion: SB in ciphertext = TH in plaintext
Hypothesis H1: S = T
H2: B = H
H3: R = E

```
THE TH      H    E T           E T    EE        T
SBR SB CTU DB CKERVS FCGG WTTCXR SFH FRRJD YTHE SHUWI
```

Comments:

At this point we have no way of telling whether A1 and A2, H1 to H3, are right (I shall explain why I prefer this term later). Above all, we have no way of testing them. We can only press on and see what happens. Obviously, the fact that we have no way of testing them is no mark against these assumptions and hypotheses. (However, if we were foolish enough to stop at this point, and insist on the truth of A1 and A2, and H1 to H3, we are likely to draw upon ourselves the complaint that we do not understand what an investigation is about.)

What if these assumptions and hypotheses are wrong?

If they are wrong, we will not be able to get very far with our analysis after this point. (Readers can easily test this for themselves). This is a general principle we observe in investigations. In an investigation, when we come to an impasse, one of the possibilities we have to consider is that we have accepted wrong hypotheses or wrong assumptions.

The formation of assumptions and hypotheses is a trial-and-error process dependent on the detection of clues. The latter—the detection of clues—is an important element in an investigation. Now in case anyone should wonder whether the existence of clues might not literally mean that we have been given the knowledge we are looking for, we would do well to remember that clues are not given; they have to be discovered. In our example, clues are obvious, and thus easy to detect. In other investigations, they may first have to be unearthed, and even then may be difficult to interpret. The discovery of a clue not usually noticed is regarded as an accomplishment. Pythagoreans take the mathematics of musical strings as a clue to the nature of reality. We are still talking about them.

What do clues do?

Some clues remind us of elements in our background knowledge. If we have known about substitution ciphers before, it is easier to 'recognize' them. In investigations, we need background knowledge. But our background knowledge is vast: it is not humanly possible to try out every single item to see if it is useful in a particular investigation. Some clues enable us to make use of our background knowledge without having to try out every single item.

(Background knowledge can be useful in different ways. Sometimes, they provide us with analogies, or approximate analogies, to the realities we are trying to uncover. When these analogies succeed, we will have extended our knowledge into new areas. The decipherment of dead languages depends on approximate analogies. The use of two different approximate analogies—waves and particles—eventually led us to a better understanding of light.)

While some clues remind us of items in our background knowledge, others enable us to take small creative leaps (Lai [1986]). If we have experienced all shades of blue except one, we can imagine the missing shade. From context, we learn new words. A new scientific instrument may suggest a new way of testing some particular hypothesis. Well defined difficulties in explaining blackbody radiation became the clue to quantum theory.

Step 2. Suspicion: GG=LL
Hypothesis H4: G=L
H5: F=W
H6: C=I

```
THE THI    HI  E T WILL    I E TW   WEE          T
SBR SBCTU DBCKERVS FCGG WTTCXR SF H FRRJD YTHE SHUWI
```

Comments:

H4 to H6 are again more or less shots in the dark; we have no indication that they are right.

Step 3. Hypothesis H7: H=O
H8: J=K
H9: D=S

THE	THI	SHI E T	WILL	I E	TWO	WEEKS		TO
SBR	SBCTU	DBCKERVS	FCGG	WTTCXR	SFH	F RRJD	YTHE	SHUWI

Comments:

H7 is the answer to a clue. The sixth word into the message contains three letters and begins with TW. This gives us a clue to H7. Since H7 is purposely constructed to complete the word TWO, the formation of this word cannot be regarded as a test for H7.

Similar to H7, H8 and H9 result from responding to a clue.

Notice one thing we have been doing. Once we have an hypothesis, we apply the hypothesis throughout the cryptogramme. This we do naturally, but it is important to ask ourselves why. The answer of course is, we are interested in uncovering the clear message. We are not only interested in hypotheses (as to how the cipher works); the clear message is just as important.

In applying hypotheses, we should notice, we are NOT necessarily testing them. We have been applying hypotheses, but we have not been testing any hypotheses so far. (See also Lai [1983]).

Now while we apply hypotheses because we are interested in uncovering the clear message, after applying them we *sometimes* develop clues. Clues are necessary if we are to continue with the analyses.

Step 4. Suspicion: Last word is TODAY.
Hypothesis 10: U=D
11: W=A
12: I=Y

THE	THI D	SHI E T	WILL	A I E	TWO	WEEKS		TODAY
SBR	SBCTU	DBCKERVS	FCGG	WTTCXR	SFH	F RRJD	YTHE	SHUWI

Comments:

Up to this point, we have not been able to test any of our hypotheses. Yet, even without testing any of them, we have a strong intuition that H1 to H9 at least are likely to have been right. Why is this the case? Is there a rational basis to our intuition?

I suggest there is and we can work it out this way. H1 to H6 are more or less shots in the dark. However, after applying them to the ciphertext, they produced clues that we did not have before. These clues led to H7 to H9. After applying H7 to H9 to the ciphertext, we had further clues, leading to H10 to H12. Now H1 to H6 could have been wrong, but if they were wrong, very likely they would have led to an impasse; they would not have produced clues which led to H7 to H9. Similarly, H7 to H9 could have been wrong, but if they were, they would not have led to H10 to H12. Thus by the time we get to H10, we would have reason to think that H1 to H9 are likely to have been right.

But why is it that if an hypothesis is wrong, it is likely to lead to an impasse? To answer this question, we have to look more carefully at what we have been doing. In our present investigation, we have a cryptogramme and we want to recover from it both the cipher and the clear message. We cannot impose on the cryptogramme some arbitrary interpretation of our own—say, through free association—and say that is the clear message. If this were our task, it could have been carried out rather easily. The interpretation we make of the cryptogramme, the clear message we are to recover from it, has to be derivable from the cryptogramme *together* with the cipher. But the cipher has not been given; we have to recover it as well. But just as we cannot arbitrarily impose an interpretation on the cryptogramme, so we cannot arbitrarily construct a cipher of our own and apply it to the cryptogramme. Such a cipher will not work. The cipher we arrive at, and the clear message, have to harmonize. This is what makes cryptanalysis difficult. In cryptanalysis, we are working under severe constraints.

How do we solve this problem? How do we recover both the cipher and the clear message?

We follow a STRATEGY. We break the problem down into manageable portions, *and we try to uncover both the cipher and the clear message at the same time, or rather in parallel.* We shall call this strategy the investigative strategy (IVS).

We have to break the problem down; it is too complex to solve in one single leap. We cannot in one single leap conjecture the whole cipher and the whole message—and then look for evidence to confirm or refute our conjectures. Since we cannot do this, we first make general assumptions concerning the hidden secrets we are trying to uncover, such as the language of the plaintext (English), the *kind* of cipher used (simple substitution). Then we try to work out the details—part by part, portion by portion. We do this by proposing hypotheses, one or a few at a time, as to how the cipher works. We apply these hypotheses to the parts of the ciphertext to which they are relevant. From the parts of the message deciphered, we, hopefully, derive clues. Clues hopefully lead to the proposal of more hypotheses. More hypotheses hopefully lead to more clues. This is how we proceed. This is how we HAVE to proceed.

In addition to working on the problem in a piecemeal fashion, we have also to uncover the cipher and the plaintext *in parallel*. We cannot hope to uncover one first and then work out the other. The cipher and the plaintext have to harmonize. By uncovering them in parallel, we are able to check the (partial) results in one process by the (partial) results in the other, and *vice versa* (see also Lai [1986]). *Because the cipher and the plaintext have to harmonize, and harmonize in a very specific way, unless the general assumptions are correct, or roughly correct, and the hypotheses, and the partial decipherments, we cannot proceed.* Here then we have the explanation why wrong hypotheses will lead to impasse. The crux of the matter is, we are here dealing with a complex process: Wrong hypotheses will lead to wrong decipherments. Wrong decipherments are either immediately detectable, or (coincidences apart) they lead to wrong clues, or no clues at all. Wrong clues will lead to further wrong hypotheses. If we have no clues, it is difficult to arrive at more hypotheses.

From the above, we see that cryptanalysis is a process in which the development of hypotheses goes hand in hand with their evaluation.

In sum, the method of evaluating hypotheses in cryptanalysis, and thus in investigations in general, is the following. If progress is halted, we suspect the presence of wrong hypotheses. If progress resumes after an hypothesis has been replaced, the replaced hypothesis is likely to be the wrong hypothesis. An hypothesis that leads to progress is likely to be right. The earlier an hypothesis is (in its involvement in the production of clues for later hypotheses), the more certain it is. The latest hypotheses are the least certain (since we do not know whether they will lead to progress). Since this method of evaluating hypotheses depends on there being progress in the investigation, I call it the method of progressive evaluation, or MPE. MPE is the method with which we have been evaluating hypotheses in our example.

To further clarify MPE, and to prevent it from being confused with other methods of evaluation, the following points may be useful.

(1) In following MPE, we are interested in whether we are making progress. Now progress in the present context does not mean arriving at an hypothesis with greater explanatory power. An hypothesis with greater explanatory power is a replacement hypothesis; it replaces one that has failed. Progress in the present context means more NEW hypotheses and more NEW decipherments. We want both. The uncovering of the cipher and the uncovering of the message have to occur in parallel.

Now it is true that the more hypotheses we have, then, together, they are able to explain more. H1 to H12 together explains more than H1 to H3. However, in following MPE we do not evaluate hypotheses as a group. Methodologically, our investigation has to be a piecemeal process. Since evaluation of hypotheses is concurrent with their development, evaluation of hypotheses has to be piecemeal also.

(2) In our example, in evaluating hypotheses by MPE, *we do not have to wait until the whole message has been deciphered* (for a real-life example of an

incomplete cryptanalysis, see Brumbaugh [1978]). Very definitely, in using this method, we are NOT saying that because a certain group of hypotheses leads to a coherent final message, that group of hypotheses must be true. This is a trap easy to fall into. We have made a point not to fall into it. In our example, it is possible that the investigation will NOT have been finished even when the whole message has been deciphered; the investigation may have to be carried on beyond that point. There are any number of reasons why this may have to be done. For example, there could be a second level concealment, a code within a code (superencipherment). Since there is this possibility, we cannot be certain that our hypotheses are true even after we have gotten a coherent message after a first level decipherment. Further investigation may require us to revise some of our hypotheses.

Now, if there is opportunity to expand an investigation, it should be welcome. If MPE is the fundamental method of evaluation, then the most important methodological advice we can give ourselves is to push ahead with investigations whenever possible. For, with progress, we not only discover more, but we become more certain of our earlier results, or have a chance to bring them closer to the truth.

(3) In cryptanalysis, we are interested in the objective truth (or should we say, truths). In those instances where we are able to have direct access to the cipher after we have arrived at a solution (as in games and in training exercises), we can even verify whether we have actually arrived at it, or have come closer to it. When we do not have direct access to the cipher, the reason we give, for thinking that we *might* have arrived at the objective truth, or at least have come closer to it, is that the harmonizing patterns we uncover—because there are two of them, and because they harmonize—could not have been our own making. When we succeed in expanding an investigation, we will arrive at more harmonizing patterns. Indeed, we are likely to arrive at an *expanding complex* of harmonizing patterns. When this happens, we have even more reason for thinking that these patterns are objective, not the results of coincidence or self-deception.

Notice that in a successful investigation, we uncover at least two harmonizing patterns.[3] Newton discovered not only the laws of motion and the theory of gravitation, he also discovered a model for the solar system. In this model, heavenly bodies, which we see as spots of light or bright discs, are *material*, not made from the fifth essence.[4]

---

[3] We do not justify one pattern by the other, nor *vice versa*. Sometimes people do both and arrive at the mistaken conclusion that even science is a form of ideology; that its reasoning is self-contained.

[4] People sometimes become so familiar with a code that they speak and read in code, without the need for translation. Scientists can become so familiar with theories that they see voltage drops and solutions turning alkaline. We are so familiar nowadays with heavenly bodies being material that we sometimes forget that this way of deciphering them was the result of the same investigation that produced Newton's theories.

I have been using the terms 'right' and 'wrong' in connection with hypotheses. The reason is the following. For us to evaluate hypotheses using MPE, the investigation has to keep advancing. Now, with some investigations, for them to advance, the hypotheses do not have to be true; they could be just good approximations; and not just good approximations quantitatively, but even be just good approximations ontologically (hence the possibility of using analogies in the formulation of hypotheses). The use of approximations, provided they are good enough, will not hurt the investigation. As the investigation advances; as it uncovers more (as we have more hypotheses and more decipherments); approximations can be improved upon, bringing them ever closer to the truth. There is no mystery in this. It is a matter of taking advantage of an expanded context. It is easier to detect subtle mistakes when reading a long passage than when reading a short one. With the increase understanding that comes with Newtonian theory, we can turn back and correct Kepler. Now since a good approximation is in a strict sense false, it may not be of any great concern to be told during an investigation that a certain hypothesis is false. It is more important to find out whether it is right, that is, whether it is close enough to the truth to help advance the investigation. In investigations, we are concerned with truth in the long term. In the short term, rightness suffices.

After this long comment, let us now go on to Step 5.

Step 5. Suspicion: Second last word is from.
Hypothesis H13: Y=F
H14: T=R
H15: H=O
H16: E=M

```
THE THIRD SHI  ME T  WILL ARRI   E TWO  WEEKS FROM TODAY
SBR SBCTU DBCKERVS  FCGG WTTCXR SFH   FRRJD YTHE  SHUWI
```

Comments:

Before step 5, the only method we have used to evaluate hypotheses is MPE. Here we put forward H14. For H14 there is an empirical test. It is made by applying the hypothesis to the second word. Before H14 is applied, the second word appears partially deciphered thus.

```
THI D
SBCTU
```

Since this word contains the ciphertext letter T, if H14 is correct, we should, after translating T, have a proper English word. If we do not, H14 is likely to be wrong. Since after applying H14, we do have a proper English word, H14, we conclude, is likely to be right. So, here, we have a test for H14.

Notice that we have a test for H14 only because all the other letters in the second word can be assumed known. If there were additional unknown letters, besides the one to which H14 is relevant, applying H14 to the second word would not have served as a test for H14. This is how we normally test an hypothesis. We test an hypothesis by applying it to a relevant situation in which all elements can be assumed known, with the exception of those to which the hypothesis is relevant. When we do this, we can tell whether the hypothesis is right by looking at the outcome, to see if it is as predicted. Since this kind of situation cannot be found so far except for H14, only H14 can be tested. However, as we have seen, the fact that the other hypotheses have not been tested does not mean that we have not been evaluating them. In fact we have. We have been evaluating them by MPE. And we could have gone on evaluating hypotheses this way—until we finish deciphering the whole message. Tests are inessential! The test for H14 is something of a bonus. Nice to have, but inessential.

Why, if tests are inessential, are they nice to have?

A moment's reflection will tell us that they are nice to have because they provide us with quick doublechecks, thus reducing the chance of major errors. We propose hypotheses based on clues. But we could have read the clues wrong. Hence it is useful to have doublechecks.

The presence of major errors may not matter much in simple investigations. If the investigation comes to a stop, it is relatively easy to retrace our steps. However, in complex investigations, once a major error has burrowed itself, it may be difficult to unearth it later on. In complex investigations, it is much more difficult to retrace our steps. Thus in complex investigations, a quick doublecheck on an hypothesis is highly desirable. By securing as much as is possible steps we have taken, quick doublechecks help cut down confusion. (In complex investigations, when a large number of mistakes have been made, sometimes it is easier to redo a portion of the investigation than to try to retrace our steps. The Copernican Revolution is one such event.)

However, while tests are nice to have, the fact remains, they are but auxiliary devices. They provide some assurance that we have not made major errors, *but this assurance is conditional,* and thus tentative: it has to be confirmed in the longer term by MPE. It is conditional because in every test, there is a major assumption. In every test situation, we assume that all the auxiliary hypotheses are true. (The test on H14 depended on H1, H2, H6 and H10 as auxiliary hypotheses). Only on this assumption can the test provide us with a quick doublecheck. But this assumption, which we may call the test assumption, can never be fully discharged: we can never be certain that any hypothesis is true. This is brought home to us on those occasions when, at a later stage in an investigation, we find that an earlier hypothesis has to be revised.

However, while the test assumption can never be fully discharged, it can be progressively discharged: we can become progressively more certain that the auxiliary hypothesis are right as the investigation advances. When the hypothesis tested leads to further clues which lead to further hypotheses, we become more certain not only of the hypothesis tested, but also of the auxiliary hypotheses used in its test. In our example in cryptanalysis, should we become more certain of H14—which will come about when progress in the investigation continues, we will become more certain of the auxiliary hypotheses, H1, H2, H6 and H10 as well. Thus, to confirm the verdict of a test, we have to wait for progress. This means, we can never escape from relying on MPE. The latter is fundamental. Tests are but auxiliary devices. They can conditionally rule out major errors (conditionally refute hypotheses) but they cannot greatly increase the certainty we have that certain hypotheses are right. Such certainty can only increase in the long term through progress.

In this explanation just given, there are two points that may benefit from further clarification.

(1) We were saying that tests can conditionally refute hypotheses. By this we do not mean to imply that we can defend an hypothesis come what may, as the Duhem-Quine thesis contends. If we defend the wrong hypothesis, we pay the price of not being able to push ahead with the investigation (see also Lai [1984]). It is true that we cannot tell from the result of a test which hypothesis is definitely false. But this does not mean that we can only refute hypotheses as a group. The reason is, tests are only auxiliary devices; the central method of evaluating hypotheses is MPE. Ultimately which hypothesis has to be rejected as false is determined by MPE, that is, by seeing the replacement of which hypothesis leads to progress. Newtonian theory was conditionally refuted by the Michelson and Morley experiment. This refutation became more certain when Newtonian theory was superseded by Einsteinian theory.

In science, in the practice of designing tests for theories, we have an indication that we do not try to refute theories as a group. In designing tests, we choose when we can auxiliary hypotheses of which we are comparatively more certain. Theories incorporated in 'well-tested' scientific instruments are theories of this sort. That we are able to distinguish the more certain from the less certain implies that we have a method of gauging certainty. This method is MPE.

(2) Some may think that evaluating hypotheses by MPE is the same as testing them; there is, they may say, no difference between the two methods of evaluating hypotheses. Now if this were so, there would be no point in saying which is central and which is auxiliary. But it is not so. Relying on MPE and relying on empirical tests are two distinct ways of evaluating hypotheses. In an empirical test, we deduce from the hypothesis under test (together with auxiliary hypotheses and statements describing initial conditions) some 'prediction'. If the prediction is false, we infer, under the test assumption, that

the hypothesis is false. In MPE, when we evaluate an hypothesis, we do not deduce from it a prediction. In MPE, we are normally interested in finding out whether an hypothesis is right. To do this, we try to see whether the hypothesis leads to progress. We cannot predict that an hypothesis will lead to progress. An hypothesis could be right (i.e. true, or a good approximation) without leading to progress. To see whether an hypothesis leads to progress, we apply the hypothesis to see whether it develops new clues (reveal new problems; see again comments under Step 3). We then try to see whether these new clues lead to more hypotheses which lead to more new clues, and so on.

With MPE, we suspect the presence of seriously defective, false, hypotheses when an investigation reaches an impasse. In MPE, to see whether an hypothesis is false, we do not deduce from the hypothesis in question a prediction. Instead, we replace it by another hypothesis and see whether after the replacement, the investigation advances again. When it does, the replaced hypothesis, we conclude, is false.

From what has been said, we can see that an empirical test depends on a direct appeal to the logical relation between premise and conclusion. When the conclusion is false, at least one of the premises must be false. MPE, on the other hand, is integral with the investigative strategy (IVS). We need an understanding of this strategy to understand why MPE works.[5] Tests are auxiliary, dispensable, devices within IVS; the central method of evaluation is MPE.

In science, there is recognition that empirical tests are only auxiliary devices. Real experiments can sometimes be replaced by thought experiments. Whether Galileo actually drop two canon balls from the Tower of Pisa, or whether he merely carried out a thought experiment, is, from the standpoint of MPE, quite immaterial; his theory of free fall provided clues to the theory of universal gravitation.

Important discoveries in science are sometimes 'based on' (tested by) crude experiments. For example, Dalton's atomic theory. But this does not matter overmuch. If the discovery leads to other discoveries, all is well—for the moment; refinements can be attempted later.

We have not finished deciphering our message. However, now that I have had opportunity to make my point, and to explain MPE and IVS, finishing the decipherment is not all that important. I leave it to my readers.

Memorial University of Newfoundland,
St. Johns, Newfoundland
Canada

---

[5] It may be useful to notice here that IVS is not any form of logic, deductive or otherwise; it is an answer to a *practical* problem, just like the airfoil. With the airfoil, we are able to fly without growing wings. With IVS we are able to uncover knowledge without direct access.

REFERENCES

BRUMBAUGH, R. S. [1978]: The Most Mysterious Manuscript: the Voynich 'Roger Bacon' Cipher Manuscript. Carbondale: Southern Illinois University Press.

HATTIANGADI, J. N. [1983]: 'To Save Fallibilism'. *Mind*, 92, 407–409.

KAHN, D. [1967]: The Codebreakers. New York: Macmillan.

KUHN, T. [1962]: The Structure of Scientific Revolutions. Chicago: University of Chicago Press.

LAI, T. [1983]: 'Why Experiments?' Abstracts, 7th International Congress of Logic, Methodology and Philosophy of Science, vol. 3.

LAI, T. [1984]: 'The Philosophical Importance of "Technically Good" Experiments'. *British Journal for the Philosophy of Science*, 35, 156–159.

LAI, T. [1986]: 'Creativity in Experimental Science'. Essays on Creativity and Science, ed. by Diana MacIntyre DeLuca. Honolulu: Hawaii Council of Teachers of English.

POPPER, K. [1959]: The Logic of Scientific Discovery. Harper Torchbooks 1968.

POPPER, K. [1962]: Conjectures and Refutations. Harper Torchbooks 1968.

Brit. J. Phil. Sci. 39 (1988), 225–232 *Printed in Great Britain*

Realism and the Collapse of the Wave-Packet

HENRY KRIPS

ABSTRACT

Cartwright's argument for the collapse of the wave-packet is criticised. This enables us to rescue realism from the pressure put on it by Cartwright in the context of Quantum Theory.

1 *Cartwright's argument for Collapse of the Wave-Packet*
2 *Explaining Cartwright's correlation*

I CARTWRIGHT'S ARGUMENT FOR (CWP)

In her [1983] Cartwright argues for the collapse of the wave-packet (CWP) as follows. She points out that for ideal conservative measurements (in which the quantity measured is conserved) the results of immediate repeat measurements always agree with the results of the initial measurements. But, her argument continues, the only way to explain such a correlation between the initial and repeat mvs (measured values) seems to be via the supposition that those measurements are 'state-preparations' (to use the physicists' terminology), that is they prepare the measured systems so that there is unit probability that the measured quantities have values equal to the mvs, that is they prepare the measured systems to be in eigenfunctions of the measured quantities corresponding to eigenvalues which are equal to the measured values. And the latter supposition is of course just (CWP). Thus (CWP) is argued for by claiming that (CWP) is a necessary adjunct to QT (Quantum Theory) if QT is to be able to explain the phenomenon of measurement, and in particular if QT is to explain the correlation between the results of successive (conservative) measurements.

My counter to this argument will be two-fold. I will show that (a) an alternative explanation of the relevant correlation is available; and that (b) QT is inconsistent with (CWP). These two moves together not only undermine Cartwright's support for (CWP) (since her supporting argument depends on (CWP) being the *only* explanation for the correlation), but also (I claim) provides good reason to reject (CWP) given the strong support that there is for QT.

By contrast Cartwright would agree with (b) but would also preserve the idea that (CWP) is necessary for QT. She would then argue that this coexistence of contradictories within the body of QT forces us to give up realism for QT. (See her [1983] p. 200). By putting forward (a), I rescue realism from this particular threat.

But first I shall demonstrate (b). Consider the Stern-Gerlach experiment. A stream of silver atoms passes in the x-direction between the poles of a magnet aligned North-South along the y-direction. It is well known that there is probability 1 that the atoms are deflected by the field positively/negatively in the z-direction, depending on whether they are in an eigenstate of the y-component of spin for eigenvalue $1/2$ or $-1/2$ respectively. That is if we let the atoms be initially in an eigenstate of z-displacement for eigenvalue 0, as described by the function F_{z0}, and if we let t and t' be the times for entering and leaving the magnetic field, then

$$U_{t't} f_{y+} \times F_{z0} = f_{y+} \times F_{z+} \tag{i}$$

$$U_{t't} f_{y-} \times F_{z0} = f_{y-} \times F_{z-} \tag{ii}$$

where $U_{t't}$ is the operator (the Schrödinger propagator) which transforms states at t into states at t', f_{y+} and f_{y-} are eigenstates of the y-spin for eigenvalues $1/2$ and $-1/2$ respectively, and F_{z+} and F_{z-} are eigenstates of the coarse-grained z-displacement corresponding to positive and negative values of z-displacement respectively. (Note that f_{y+} appears on both sides of the first equation, as does f_{y-} for the second equation, reflecting that this is a conservative measurement). In addition we will suppose that these positive and negative values of z-displacement are macroscopically significant (by letting the field be strong enough).

Now let the atoms' initial state function be a superposition of f_{y+} and f_{y-}; in particular let it be

$$(c_1 f_{y+} + c_2 f_{y-}) \times F_{z0}$$

Then QT tells us (from the linearity of the equations of motion) that at t' the ions' state-function is

$$\Phi = c_1 f_{y+} \times F_{z+} + c_2 f_{y-} \times F_{z-} \tag{iii}$$

Crucial point. I claim that a y-spin measurement has been successfully completed by the time t', even though no observer actually makes an observation of the location of the atoms at the time t'. I justify this claim by pointing out that by the time t' a correlation has been established between on the one hand macroscopically distinct possible states of the atoms (namely F_{z+} and F_{z-}) and on the other hand their y-spin eigenstates (f_{z+} and f_{z-}). The fact that there has been no actual observation (however coarse) of the z-displacement of the atoms is irrelevant. After all (I claim) a measurement merely requires the production of a macroscopically significant trace, not that the trace in question has actually been registered in some observer's consciousness (although it may of course go on to be just that). This view of measurement has been argued for by Ludwig [1954] for example.

We can now derive a contradiction with (CWP). There are two ways of doing this depending on how we interpret (CWP). One way to interpret (CWP) is as the requirement that the state of *each* atom at t' is either $f_{y+} \times F_{z+}$ or $f_{y-} \times F_{z-}$. This interpretation of (CWP) presupposes interpreting state-functions as descriptions of the states of *individual* systems. But then the contradiction with (iii) is immediate since we must also take Φ to describe *each* atom's state, and $\Phi \neq f_{y+} \times F_{z+}$ or $f_{y-} \times F_{z-}$.

Alternatively we can interpret (CWP) more weakly, and combine it with the Born statistical interpretation, to give the requirement that the atoms at t' are in a mixed state, for which there is a probability $|c_1|^2$ associated with the state $f_{y+} \times F_{z+}$ and probability $|c_2|^2$ associated with the state $f_{y-} \times F_{z-}$. Formally (and this is all we need for current purposes) this means that the atoms at t' have a density operator:

$$\mathrm{W} = |c_1|^2 \mathrm{P}(f_{y+} \times F_{z+}) + |c_2|^2 \mathrm{P}(f_{y-} \times F_{z-}) \qquad \text{(iv)}$$

The latter interpretation of (CWP) has the advantage that it can be used irrespective of whether we take the states of QT as applying to individual systems or to ensembles of systems. Note however, that the latter interpretation of (CWP) entails the former interpretation of (CWP) if we assume the so-called Ignorance Interpretation of Mixtures (discussed by van Frassen on p. 171 of his [1980]).

But (iv) also contradicts (iii). For example if we measure $s_z \times P_z$ (the product of z-component spin and momentum) at t' then it is easy to show that the expected measured value differs, depending on whether at t' the state of the silver atoms is given by (iii), as demanded by QT, or by (iv), as demanded by (CWP). (Note that this contradiction only shows up by measuring quantities which commute with neither s_y or z and are joint, in the sense that they can only be defined on the product of the spin and position eigenfunction spaces). And therefore, on this second interpretation as well, (CWP) is inconsistent with QT.

Note the special role played in the above proof by the assumption that measurements can be observer-free. It is only because of this assumption that we can take the Stern-Gerlach interaction by itself (i.e. just the interaction of the atoms with the magnetic field) to constitute a measurement process, and hence are able to put it forward as a measurement process for which we know the detailed dynamics. If we needed to add an interaction with an observer to the Stern-Gerlach interaction in order for it to be counted as a measurement, then the measurement as a whole would pass outside of the range of phenomena for which we know the detailed dynamics and hence the above contradiction would no longer be derivable.

2 EXPLAINING CARTWRIGHT'S CORRELATION

We now come to our central problem. How do we explain Cartwright's correlation (between successive measured values) without resorting to (CWP)? The correlation in question can easily be demonstrated in the case of the Stern-Gerlach experiment by the following extension of the Stern-Gerlach set-up. Place a second magnet M_2 parallel to but behind the first magnet M_1, so that s_y is measured for all atoms emerging from M_1 by deflecting them either positively (in which case s_y is measured to have value $1/2$) or negatively (in which case s_y is measured to have value $-1/2$) by macroscopically significant amounts in the z direction. The Cartwright correlation then requires that all (and only) those atoms which are deflected z positively by the first measurement (with M_1) are further deflected z positively by the second measurement (with M_2).

Clearly (CWP) could explain this correlation as follows. The second measurement registers value $1/2$ for just those atoms which the first measurement projected (collapsed) into the y-spin eigenstate for eigenvalue $1/2$. But those atoms (according to (CWP)) are just the atoms for which the first measurement registers the value $1/2$. The question we now face however, is how else can we explain this correlation, if not by (CWP)? My suggestion is to first interpret the density operator for a quantum system at a particular time as a state-description for the *individual* quantum system at that time (rather than for an ensemble of similar quantum systems). This implies *inter alia* some sort of propensity theory for the relevant probabilities in order to render them both objective and single-case. Moreover I shall take it to be the relevant density operator for S at t which tells us whether a particular quantity Q for S has a determinate value at t. In particular Q has a determinate value for S at t iff $W(S,t)$, the density operator for S at t on the smallest Hilbert space for S for which Q is represented by an operator, is diagonal in a complete orthonormal set of eigenfunctions of the operator representing Q. (Note that we here allow that the one system S is associated with several different Hilbert spaces. For example in the case of a silver atom there is a spin space and a configuration space, as well as the direct product of these spaces.) This criterion for determinancy is not just *ad hoc*, but can be defended by its broader role in interpreting QT—see Krips [1987].

It then follows trivially from (iii) that it is determinate for each atom at t' whether it is in the $z+$ or $z-$ region. This is because taking the trace of $\mathbf{P}(\Phi)$ over the Hilbert space to which f_{y+} and f_{y-} belong produces the density operator:

$$|c_1|^2\mathbf{P}(F_{z+})+|c_2|^2\mathbf{P}(F_{z-})$$

which is diagonal in the coarse-grained z-displacement eigenfunctions, and because the z-displacement of the ion is defined on no smaller Hilbert space than the space to which F_{z+} and F_{z-} belong. Similarly we can show that the y-spin for each atom at t' is determinate, although what value of y-spin each has is only statistically determined.

In addition we assume that if the relevant $W(S,t)$ is diagonal in a complete orthonormal set of eigenfunctions of the operator representing Q then not only is Q determinate in value at t but also the various possible events of Q taking the value q_i in S at t, for the various q_i which are the corresponding eigenvalues of the operator representing Q, are mutually exclusive and exhaustive (in the sense that the sum of the probabilities associated with these various possible events is 1, and the joint probability of any two of those events is 0). In particular we assue that if $W(S,t) = \Sigma p_i \, P(f_i)$ then there is probability p_i that Q has value q_i in S at t for each i. Thus not only does the y-spin for each atom at t' have a determinate value, but also the possible events $(s_y+$ and (s_y-) are mutually exclusive and exhaustive, where '(s_y+)' here denotes the event of the y-spin for one of the atoms—say the atom S—having value 1/2 at t', and '(s_y-)' denotes the event of the y-spin of S having value $-1/2$ at t'.

We also assume that for any ideal measurement of the quantity Q the various possible events of the measuring apparatus registering the value q_i at the end of the measurement, for the various q_i which are the eigenvalues of any operator representing Q, are mutually exclusive and exhaustive. In particular the events $(z+)_1$ and $(z-)_1$ are mutually exclusive and exhaustive, where '$(z+)_1$' denotes the event of the z-displacement of S being positive at t', and this constitutes the event that the measurement of the y-spin of S which begins at t registers the value 1/2 by the time it finishes at t' (and similarly for $(z-)_1$). Similarly $(z+)_2$ and $(z-)_2$ must be mutually exclusive and exhaustive.

And finally we assume that ideal measurements are conditionally statistically passive, which means that for any r the conditional probability of the measured value of Q being r for a measurement of Q given that Q possesses the value r, is 1. In particular this entails that

Prob $((z-)_2/(s_y-)) = 1$, where the probability here is (v)
conditional on the second measurement at t' taking place.

The Cartwright correlation between successive measured values can now be explained as follows. The *explanandum*, together with the above assumptions, can easily be shown (see Appendix) to imply that, under the condition of both measurement at t and at t' taking place, any atom having a positive z-displacement at t' is perfectly correlated with its having a y-spin value 1/2 at t'. And the latter correlation between possessed values (rather than measured values) can then be used for straightforwardly explaining the Cartwright correlation.

The explanation is of a perfectly standard causal variety: the initial measurement by M_1 sets up a correlation between (s_y+) and $(z+)_1$ (note that this correlation is weaker than (CWP)); the measurement by M_2 then correlates (s_y+) and $(z+)_2$ (since it is passive); and hence $(z+)_1$ and $(z+)_2$ are correlated. Thus we have the not unfamiliar pattern:

Explanandum justifies *explanans*
Explanans causally explains *explanandum*

(See Scriven [1969] for a discussion of this explanatory pattern). In short we have fulfilled our programme of giving an alternative explanation of Cartwright's correlation without invoking (CWP), and thereby defused Cartwright's attack on realism in the context of QT.

It is important to note the following feature of the interpretative principles for the density operator which I have introduced above. They do *not* licence the inference:

(Inf) Q has value q_i in S at t with probability p_i

$\therefore$ S at t is in the pure state f_i with probability p_i

where f_i is the unique eigenfunction of Q for eigenvalue q_i. Instead they licence a sort of converse of this inference, namely from $\mathbf{W}(S,t)=\Sigma p_i\mathbf{P}(f_i)$ to Q has value q_i with probability p_i for all i. Indeed, as we shall now show, the above interpretative principles are actually inconsistent with (Inf), as is the Ignorance Interpretation of Mixtures (according to which $\mathbf{W}(S,t)=\Sigma p_i\mathbf{P}(f_i)$ iff there is a probability p_i that S at t is in the pure state f_i for all i).

This inconsistency can be seen to arise as follows. Consider a combined system $S+S'$ in the entangled state $\Sigma c_i f_i \times g_i$ at t, where $\{f_i\}$ are eigenvectors of a quantity Q represented by the non-degenerate operator $\mathbf{Q}$ on some Hilbert space $H(S)$ for S, and similarly the $\{g_j\}$ are eigenvectors for a quantity Q' for S'. Then $\mathbf{W}(S,t)$ on $H(S)$ is $\Sigma|c_i|^2\mathbf{P}(f_i)$ and $\mathbf{W}(S',t)$ on $H(S')$ is $\Sigma|c_i|^2\mathbf{P}(g_i)$; and hence, by combining our density operator criterion for quantities to have determinate values with (Inf) (or indeed the Ignorance Interpretation of Mixtures), we derive both that S at t is in a pure-state f_i for some i, and that S' at t is in a pure state g_j, for some i,j. And this in turn means that $S+S'$ at t is in a pure state $f_i \times g_j$ for some i,j, which contradicts our initial hypothesis that $S+S'$ at t is in the pure state $\Sigma c_i f_i \times g_i$. What I suggest is that we avoid this contradiction by rejecting both (Inf) (thus breaking the eigenvector-eigenvalue connection in one of its forms) as well as the Ignorance Interpretation of Mixtures.

To simply reject (Inf) and the Ignorance Interpretation in this way is not satisfactory however; we must also suggest a viable alternative interpretation for mixtures. This task, on which we have made some inroads above in suggesting principles for interpreting the density operator, is completed in Krips [1987]. We can only make two brief points about this interpretation here. First, because of the density operator criterion for physical quantities to have determinate values which we gave above, our interpretation is committed to allowing that, in certain situations (namely if $W(S,t)$ is the *identity* operator on $H(S)$), *every* physical quantity for S at t has a determinate value (including both members of any incommensurable pair of physical quantities). This consequence forces us to de-Ockhamize the representation for physical quantities in QT (as in van Frassen [1973]); but also (via the Bell-Stapp-Eberhard argument, as discussed in Clauser and Shimony [1978]) we are led to give up the one-to-one connection between counterfactually construed measured values and possessed values for physical quantities—thereby avoiding Bell's inequality (although without embracing non-locality). But discussion of these issues takes us well beyond what we can discuss here.

Department of History and Philosophy of Science,
University of Melbourne, Victoria, Australia

BIBLIOGRAPHY

CARTWRIGHT, N. [1983]: *How the Laws of Physics Lie.* CUP.
CLAUSER, J., SHIMONY, A. [1978]: *Rep. Progr. Phys.*, 41, pp. 1881–1926.
KRIPS, H. [1987]: *The Metaphysics of Quantum Theory.* Clarendon.
LUDWIG, G. [1954]: *Die Grundlagen der Quantenmechanik.* Springer.
SCRIVEN, M. [1962]: *Minnesota Studies in the Philosophy of Science,* 3, ed. Maxwell G. and Feigl, H. Minnesota University Press.
VAN FRASSEN, B. [1973]: In *Contemporary Research in the Foundations and Philosophy of Quantum Theory,* 2, ed. Hooker, C. Reidel.
VAN FRASSEN, B. [1980]: *The Scientific Image.* Clarendon.

APPENDIX

The *explanandum* consists of the two equations:

$$P((z-)_2/(z-)_1)=1 \tag{vi}$$

$$P((z+)_2/(z+)_1)=1 \tag{vi}$$

where the probabilities here are conditionalised on both measurements (the first at t the second at t') taking place.

Now from (v) in the main body of the text it follows that

$$P((z-)_2\cdot(s_y-))=P(s_y-)$$

But, since $(z-)_2$ and $(z+)_2$ are mutually exclusive and exhaustive, $P(s_y-)=P((s_y-)\cdot(z-)_2)+P((s_y-)\cdot(z+)_2)$; and hence $P((s_y-)\cdot(z+)_2)=0$. Moreover, since (s_y-) and (s_y+) are mutually exclusive and exhaustive, $P((z+)_2)=P((z+)_2\cdot(s_y-))+P((z+)_2\cdot(s_y+))$; and hence $P((z+)_2)=P((z+)_2\cdot(s_y+))$, which means that

$$P((s_y+)/(z+)_2)=1 \tag{vii}$$

the probability here being conditional on the measurement at t' taking place.

But from (vii) and (vi)′ it follows by a standard inference of probability theory[1] that $P((s_y+)/(z+)_1)=1$, where the latter probability is conditional on both the measurement at t and at t' taking place, which is precisely one half of the *explanans*. The other half of the *explanans*, namely that $P((s_y-)/(z-)_1)=1$, can then be similarly inferred from the other half of the *explanandum*, namely from (vi), together with another aspect of the passivity condition:

$$P((z+)_2/(s_y+))=1 \qquad \text{(v)}'$$

---

[1] The relevant inference principle here is:

(a) $P(A/BM)=1$ (b) $P(A'/A)=1$
$\therefore P(A'/BM)=1$

We can prove this as follows. From (a), $P(\bar{A}BM)=0$, and, from (b), $P(\bar{A}'A)=0$. But $P(\bar{A}'BM)=P(\bar{A}'ABM)+P(\bar{A}'\bar{A}BM)$, where the right hand side terms, as we have just seen, must be 0. Hence $P(\bar{A}'BM)=0$ too; which, since $P(BM)=P(A'BM)+P(\bar{A}'BM)$ implies that $P(A'BM)=P(BM)$, so that $P(A'/BM)=1$.

Brit. J. Phil. Sci. 39 (1988), 233–246 Printed in Great Britain

Quantum Physics and the Identity of Indiscernibles*

STEVEN FRENCH AND MICHAEL REDHEAD†

ABSTRACT

This paper is concerned with the question of whether atomic particles of the same species, i.e. with the same intrinsic state-independent properties of mass, spin, electric charge, etc, violate the Leibnizian Principle of the Identity of Indiscernibles, in the sense that, while there is more than one of them, their state-dependent properties may also all be the same. The answer depends on what exactly the state-dependent properties of atomic particles are taken to be. On the plausible interpretation that these should comprise all monadic and relational properties that can be expressed in terms of physical magnitudes associated with self-adjoint operators that can be defined for the individual particles, then the weakest form of the Principle is shown to be violated for bosons, fermions and higher-order paraparticles, treated in first quantization.

I INTRODUCTION

In the philosophical literature the notions of *identity, individuality* and *indistinguishability* have received a great deal of attention. In this paper we shall discuss how these notions have been treated by quantum physicists, and discuss the relevance of this to the traditional philosophical arguments about these matters.

Identity is a relation that may exist between items that figure in a particular area of discourse. Item *a* is *identical* with item *b*, symbolically $a=b$, means informally that there are not in reality *two* distinct items at all, but only *one* item, which may be referred to indifferently as *a* or *b*. We use 'item' here as a neutral word that comprehends both *universals* and *particulars*. Particulars that *exist* in the physical world we shall refer to as *physical individuals* or just individuals for short. The existence of such individuals we take to imply their persistence through time. They are *continuants*. Physical individuals are to be distinguished then from other sorts of particular such as events or states of affairs that may be said to *occur* or *obtain* rather than to exist. In the philosophical literature individuals which have more or less well-defined spatial locations are often referred to as 'things'. What confers particularity, or individuality as we shall call it, on physical individuals? This raises the fundamental problem of how the particular is related to the universal, of how an individual is related to its attributes. Is it possible for two (non-identical) individuals to have all the same attributes in common, that is to say to be *indistinguishable* (indiscernible in traditional philosophical terminology)? Leibniz famously claimed that this was *not* possible, his Principle of the Identity of Indiscernibles (PII). In terms of second-order logic with equality PII states

---

* Some of the arguments in this paper appeared in a thesis submitted by one of us (S.F.) in partial fulfilment of the requirements for the PhD degree of the University of London, in 1984, entitled 'Identity and Individuality in Classical and Quantum Physics'.

† Present address: Department of History and Philosophy of Science, University of Cambridge, Free School Lane, Cambridge CB2 3RH

$$\forall F\ (F(a) \longleftrightarrow F(b)) \rightarrow a = b \qquad *$$

where *a* and *b* are any two individual constants and F is a predicate variable ranging over the possible attributes of these individuals.

* should be contrasted with the Principle of the Indiscernability of Identicals.

$$a = b \rightarrow \forall F\ (F(a) \longleftrightarrow F(b)) \qquad **$$

**, is uncontroversial, provided the attributes do not involve intensional contexts. If *two* individuals are identical, so there is in reality just one individual, then there can *only* be *one* set of truly predicated attributes. But * has led to a great deal of argument. What sort of attributes should be included in the range of the variable F? Should we include relations (non-monadic properties)?[1] If we include the attribute 'being indentical with *a*', which is certainly true of *a*, then *, is a *theorem* of second-order logic. But suppose we rule out trivializations of PII of this sort, on the grounds that such an attribute is not a genuine *monadic* property, but expresses a *relation* of *a* to itself which relation (of identity) *is* also satisfied by *b*, then we can still distinguish a *weak* and *strong* version of PII.

Weak version: F includes properties of spatial location.
Strong version: F excludes properties of spatial location.

Leibniz himself apparently subscribed to the strong version of PII. Hence his interest in searching for indistinguishable leaves in the Herrenhausen Gardens in Hanover.[2]. If we subscribe to PII in the weak version this raises important questions concerning what we are to understand by spatial location. On a relational theory of space a circularity threatens. Individuation of material objects—leaves, tables, chairs—involves specifying their location in space. But this location involves their relations with other physical individuals comprising a reference frame. Unless the world is sufficently asymmetrical that each object always bears a set of qualitative spatial relations to all other objects which it shares with no other object, the individuation of the reference frame cannot itself rely on spatial relations. Absolute theories of space avoid this possible circularity but at the expense of requiring an account of what it is that confers individuality on the points of space.

---

[1] There is much discussion in the literature as to whether a clear-cut distinction can be made between relational and monadic properties. For useful comments see Hoy [1984].

[2] There is no real concensus about Leibniz's own views on the status of PII. The *locus classicus* is generally held to be Leibniz (1961) Book II Chapter XXVII. For an influential critical discussion see in particular Ishiguro [1976].

PII seeks to reduce the particular to the universal, the individual to a bundle of properties or attributes. But many philosophers have argued that such a reduction is not possible, that individuals involve something over and above their attributes, that confers individuation in an essential and unanalyzable way. This is to go the way of the Lockean substratum, the unknowable 'something' that attributes 'attach' to.

If an individual acquires its individuality by something that transcends its attributes we shall say that it exhibits *transcendental individuality*, TI for short.[3] Modern analytical philosophy has not taken kindly to TI. Physical individuals are usually regarded as being individuated by their location in space despite the problems just referred to (which may threaten to reintroduce TI for points of space!), and the problem of *reidentification*, of the grounds for claiming that an individual *b* at time t is the same individual as the individual *a* at an earlier time t′, involves as necessary condition the spatio-temporal continuity of the trajectory joining the location of *a* at time t′ and the location of *b* at time t. The proponents of TI might agree that spatio-temporal continuity (STC) is what allows us to *infer* a reidentification across an interval of time, but they would claim that it is the persistent TI, that, ontologically speaking, confers the reidentifiability.

2 THE PROBLEM OF QUANTUM STATISTICS

In classical nineteenth century atomic physics the strongest version of PII, with attributes restricted to monadic intrinsic properties, i.e. 'internal' properties independent of spatial location, is clearly false for the atoms. The weak version is however true, if we allow an impenetrability assumption (IA) to the effect that two distinct atoms can never occupy the same location in space. We want to begin our discussion of quantum physics by considering an argument to the effect that quantum particles cannot be regarded as individuals at all. If this were the case, the problem of how they are individuated simply would not arise. PII would not be either true or false but simply inapplicable. The argument runs like this. Consider the problem of distributing two quantum particles having the same intrinsic properties such as mass, spin and charge,[4] but initially supposed to be individuals and labelled 1 and 2, among two possible pure quantum states $|a^r>$ and $|a^s>$, which we may suppose to be eigenstates of some maximal observable A for either particle with eigenvalues a^r and a^s as indicated by the notation for the states. (In what follows we assume for simplicity that all observables under discussion have a discrete spectrum). By analogy with the situation in classical physics we might suppose that there are four possibilities:

---

[3] This terminology is due to Post [1963].

(1) Both particles are in the state $|a^r>$
(2) Both particles are in the state $|a^s>$
(3) Particle 1 is in state $|a^r>$ and particle 2 in state $|a^s>$
(4) Particle 1 is in state $|a^s>$ and particle 2 in state $|a^r>$

Now in classical statistical mechanics arrangements 3 and 4 would be counted as distinct and given equal weight in assigning probabilities. But in quantum statistics, whether bosonic or fermionic, the arrangements 3 and 4 are counted as one and the same arrangement for the purpose of assigning weights. This is taken to show that the two *arrangements* are not only indistinguishable but are actually identical. But ontologically speaking these two arrangements are *not* identical if the two quantum particles are individuals. Hence the quantum particles cannot be individuals.

In passing we may note that this argument, while purporting to show that quantum particles fall outside the scope of PII, since they are not individuals, is also sometimes invoked to show that PII does apply to the *states of affairs* represented by the two arrangements 3 and 4.[5] What about arrangements 1 and 2? This is where bosons differ from fermions. For bosonic particles 1 and 2 are allowed arrangements to be counted with equal weight as compared with the single identified 3-cum-4 arrangement. For fermionic particles however arrangements 1 and 2 are not permitted at all. This is the famous Pauli Exclusion Principle.

We return in the next section to discuss the significance of this difference from the point of view of PII. But first we want to explain what is wrong with the argument concerning the identity of the arrangements 3 and 4. We begin by writing down the state vectors for the combined two-particle system corresponding to the arrangements 1, 2, 3 and 4. They are:

$$|a^r> \otimes |a^r> \tag{1}$$

$$|a^s> \otimes |a^s> \tag{2}$$

$$|a^r> \otimes |a^s> \tag{3}$$

---

[4] In the physics literature such particles are often referred as 'identical'. In our terminology this would mean they were one and the same particle! We shall use the term 'indistinguishable', in place of the physicist's 'identical'.

[5] This interpretation of PII is mooted in Lucas [1984] p. 131.

and

$$|a^s > \otimes |a^r > \tag{4}$$

where we use the convention that in a tensor product of two states the left-hand member refers to particle 1 and the right-hand member to particle 2.

Now it is quite true that if the quantum particles are individuals then the states (3) and (4) are not identical. But the important point to notice is that these states are *not* the ones used in discussing quantum statistical mechanics. The relevant states for that purpose are as follows:

$$|a^r > \otimes |a^r > \tag{5}$$

$$|a^s > \otimes |a^s > \tag{6}$$

$$\frac{1}{\sqrt{2}}(|a^r > \otimes |a^s > + |a^s > \otimes |a^r >) \tag{7}$$

and

$$\frac{1}{\sqrt{2}}(|a^r > \otimes |a^s > - |a^s > \otimes |a^r >) \tag{8}$$

The four states (5), (6), (7) and (8) are mutually orthogonal and span the same subspace as the states (1), (2), (3) and (4), but they are chosen so that (5), (6) and (7) are symmetric under exchange of particle labels (i.e. under exchange of left-hand and right-hand members of tensor products), while (8) is antisymmetric (changes sign) under the same operation.

Note that (5) and (6) are the same states as (1) and (2). The crucial difference is between the pairs (3) and (4) and (7) and (8).

Now (7) is no more identical with (8) than is (3) with (4). But for bosons the states are restricted to the three *symmetric* possibilities. That is why (8) gets eliminated from the counting procedure, *not* because it gets identified with (7).

Similarly for fermions the states are restricted to the antisymmetric possibilities. But in this simple example this eliminates (5), (6) and (7), so (8) alone gets counted, but again not because it gets identified with (7).

To put the matter another way, states with the wrong symmetry get eliminated because they are not *accessible* to the joint quantum system, not because there are no such states!

It should be remembered that for time-evolution under a symmetric Hamiltonian[6] the symmetry character of a state cannot change with time, so no transitions can occur between symmetric bosonic states and antisymmetric fermionic states.

The upshot of our argument is to show, not that quantum particles *must* be individuals but rather that it is *possible* for them to be individuals, despite the peculiarities of quantum statistics. It is quite true that in quantum field theory (QFT) particles are not regarded as individuals. They are simply (quantized) excitations of a field. Particle labels do not enter into the discussion at all.

---

[6] The justification for assuming that the Hamiltonian observable must be a symmetric function of the particle labels will become apparent in Section 3.

If our simple problem of counting the number of states for a two-particle system distributed over two one-particle states were transposed to quantum field theory, then for a bosonic (commuting) field there would be just *three* states corresponding to a double excitation of either state (mode) plus a single excitation of both states (modes). Similarly for a fermionic (anticommuting) field, there is only *one* state since double excitations are not allowed. So the quantum statistics comes out the way we want it to. It is also true that there are strong arguments for regarding the 'quantized excitation' view of quantum particles as the correct one.[7] However, for the purposes of this paper, we continue to envisage the possibility of treating quantum particles as individuals and proceed to discuss whether they would or would not obey PII.

3 THE INDISTINGUISHABILITY POSTULATE

What do we mean by saying that two quantum particles of the same species (characterized by their intrinsic properties) are indistinguishable?

In quantum mechanics (QM) this is expressed by the Indistinguishability Postulate (IP)

$$< P\phi|Q|P\phi > \; = \; < \phi|Q|\phi > , \forall\, Q, \forall \phi \qquad (9)$$

where $|\phi >$ is an arbitrary N-particle state and Q a posible observable on the N-fold tensor product space of states. $|P\phi >$ is an abbreviation for $P|\phi >$ where P is the unitary operator which is associated with an arbitrary permutation of the particle labels.

(9) says that it is not possible to tell by measuring the expectation value of any observable, whether the state of the system is $|\phi >$ or $|P\phi >$.

We notice that a sufficient condition for (9) to hold is that $|P\phi > \; = \; \pm \; |\phi >$ with Q *any* self-adjoint operator on the N-particle state-space. This interprets (9) as a restriction on the possible *states* for the N-particle system, allowing just the boson or fermion possibility. (Note that the choice of signs needs only to be made for transpositions, since any permutation can be represented as a product of transpositions so even permutations are always associated with the plus sign, the distinction between bosons and fermions only arising for odd permutations).

But Greenberg and Messiah ([1964]) pointed out that (9) should, on a more profound analysis, be interpreted not as a restriction on states, but as a restriction on the possible *observables* for the N-particle system. On this view (9) can easily be shown to imply $P^{-1}QP = Q$ or $QP = PQ$, so any permitted Q must commute with any permutation P. This in turn implies that Q must be a symmetric function of the particle labels. The label permutations provide effectively a set of non-Abelian superselecting operators, which can be used to resolve the state space into non-combining sectors associated with irreducible representations of the symmetric group S_N. For two particles there are only two irreducible representations of S_2, provided by states which are symmetric or antisymmetric in the particle labels. So we are back to the boson and fermion possibilities but, with more than two particles, higher-dimensional representations of the symmetric group exist, allowing for the possibility of so-called parastatistics intermediate in character between bosonic and fermionic behaviour.

---

[7] See Redhead [1983] and [1988] for further discussion of the QFT approach.

But even in the two-particle case it should be noted that the Messiah and Greenberg approach does not restrict the available states, only their accessibility in the way we described in the previous section.

Let Q now denote a possible observable (self-adjoint operator) on a single particle. Considered as possible physical magnitudes for the joint system we have two possibilities: $Q \otimes I$ for particle 1 having the property Q and $I \otimes Q$ for particle 2 having the property Q. Denote by $Q \otimes I$ by Q_1 and $I \otimes Q$ by Q_2. Then IP says, on the Messiah-Greenberg interpretation, that although Q_1 and Q_2 are self-adjoint operators on the Hilbert space for the joint system, they cannot actually be observed. The intuition here is that observing Q_1 or Q_2 would involve knowing empirically which particle was which, and this is impossible if the particles are indistinguishable.

But from the point of view of discussing PII, it seems clear that we should not restrict the discussion to attributes which can actually be observed. This would restrict the discussion to symmetric combinations such as $Q_1 + Q_2$. The *ontological* significance of PII can only be brought out by discussing whether particles 1 and 2 have the same physical attributes expressed by Q_1 and Q_2 and their associated 'actualization' probabilities, while recognizing that these attributes can never be observed.[8] This is the task we shall attempt in the next section.

4 QUANTUM INDIVIDUALS AND THE IDENTITY OF INDISCERNIBLES

We begin by discussing the case of fermions. It has been claimed in the literature that the Pauli Exclusion Principle, prohibiting two fermionic particles from being in the same quantum state, is a clear vindication of PII[9].

---

[8] We refer to 'actualization' rather than 'measurement result', to emphasize the point that they are not observable. They will, however, be produced by measurement interactions, but not in a way which makes them identifiable as contrasted with their particle label permuted variants. But note that the probability of observing some eigenvalue q^α for Q on *one or other* particle is calculated in QM as: Prob $(Q_1 = q^\alpha)$ + Prob $(Q_2 = q^\alpha)$ − Prob$(Q_1 = q^\alpha \ \& \ Q_2 = q^\alpha)$.

We are claiming here ontological significance for each individual term in this formula for an *observable* probability.

[9] See Shadmi [1978].

What is being prohibited, apparently, is that the two fermions shall have both the same intrinsic state-independent properties of mass, spin, electric charge, etc and the same state-dependent properties expressed by expectation values of all quantum-mechanical physical magnitudes. But look at the allowed state (8). It is not true in such a state that each particle is present in a different state. *Each* particle clearly 'partakes' of *both* the states $|a^r>$ and $|a^s>$ in the superposition of product states expressed in (8). So might it not appear that in the *allowed* state both particles also have the *same* state-dependent properties, which would contradict PII? Let us formulate the state-dependent properties in terms of physical magnitudes such as Q_1 and Q_2 pertaining to each particle separately as discussed in the preceding section. In orthodox interpretations of quantum mechanics the properties Q_1 and Q_2 must be interpreted not as possessed values, but as propensities to yield specified 'actualization' results in accordance with the familiar statistical algorithm for computing the associated probabilities. Denoting the fermion state (8) by $|\Psi>$ we shall be interested in comparing both monadic properties of the form $\text{Prob}^{|\Psi>}(Q_1 = q^\alpha)$ and $\text{Prob}^{|\Psi>}(Q_2 = q^\alpha)$, where the notation indicates the probability in the state $|\Psi>$ that the physical magnitude pertaining to either particle actualizes with the indicated value, and also relational properties of the form $\text{Prob}^{|\Psi>}(Q_1 = q^\alpha/Q_2 = q^\beta)$ and $\text{Prob}^{|\Psi>}(Q_2 = q^\alpha/Q_1 = q^\beta)$ which refer to the conditional probabilities of actualizing one magnitude given the actualization result for the other.

These quantities are easily computed from the joint distribution

$$\begin{aligned}
&\text{Prob}^{|\Psi>}(Q_1 = q^\alpha \,\&\, Q_2 = q^\beta) \\
&= |<q^\alpha|<q^\beta|(|\Psi>)|^2 \\
&= \tfrac{1}{2}|<q^\alpha|a^r>|^2 \cdot |<q^\beta|a^s>|^2 \\
&+ \tfrac{1}{2}|<q^\alpha|a^s>|^2 \cdot |<q^\beta|a^r>|^2 \\
&- \text{Re} <a^r|q^\alpha><q^\alpha|a^s><a^s|q^\beta><q^\beta|a^r> \qquad (10)
\end{aligned}$$

Summing this result over α and β to obtain the marginal probabilities and remembering $\sum_n |q^n><q^n| = I$, $<a^r|a^s> = 0$ and $<a^r|a^r> = <a^s|a^s> = 1$ yields immediately

$$\begin{aligned}
&\text{Prob}^{|\Psi>}(Q_1 = q^\alpha) = \text{Prob}^{|\Psi>}(Q_2 = q^\alpha) \\
&\tfrac{1}{2}|<q^\alpha|a^r>|^2 + \tfrac{1}{2}|<q^\alpha|a^s>|^2 \qquad (11)
\end{aligned}$$

Similarly we find

$$\begin{aligned}
&\text{Prob}^{|\Psi>}(Q_1 = q^\alpha/Q_2 = q^\beta) \\
&= \text{Prob}^{|\Psi>}(Q_2 = q^\alpha/Q_1 = q^\beta)
\end{aligned}$$

$$
\begin{aligned}
&= [|<q^\alpha|a^r>|^2 \cdot |<q^\beta|a^s>|^2 + |<q^\alpha|a^s>|^2 \cdot |<q^\beta|a^r>|^2 \\
&-2\mathrm{Re} <a^r|q^\alpha> \ <q^\alpha|a^s> \ <a^s|q^\beta> \ <q^\beta|a^r>]/ \\
&[|<q^\beta|a^r>|^2 + |<q^\beta|a^s>|^2] \qquad (12)
\end{aligned}
$$

The significance of (11) and (12) is that the two fermions in the state (8) do in fact have the same monadic properties and the same relational properties one to another, so the weakest form of PII which we can formulate which involves both monadic properties *and* relational properties, is violated.[10]

There are a number of comments we want to make concerning this conclusion and the way it was derived.

(1) In classical physics the state-dependent properties of a particle are completely specified by the maximally specific state description (location in phase space). Hence we can replace the question, 'Do classical particles have the same state-dependent properties?' with the question 'Do the two particles have the same maximally specific state description?'

If we try the same move in quantum mechanics we run into the problem that for a so-called 'entangled' state such as (8) there are no pure states which can be ascribed to the separate particles. (If there were such states the state of the combined system would be the tensor product of the states in question, but (8) is not of the form of a tensor product—it is a superposition of tensor products).

Now pure states in QM play the role of maximally specific states, so if we identified the relevant properties of the two particles with the pure states they are in, we would have to conclude that there is no answer to the question 'Do they have the same properties?'

A corrolary of this result is that insofar as we can speak of states for the separate particles at all we must speak of mixed states.[11] Indeed the relevant mixed states are the same for the two particles[12]—equiprobable mixtures of the states $|a^r>$ and $|a^s>$. This is, of course, the essential content of the result for the marginal probability distributions for Q_1 and Q_2. But our analysis has gone beyond that involving the (improper) mixed states of the separate particles, by considering also the relational conditional probabilities[13] given in (12).

---

[10] This result remains true if we consider also relational properties of the form $\mathrm{Prob}^{|\Psi>}$ $(Q_1 = q^\alpha / Q'_2 = q^\beta)$ which is easily seen to be equal to $\mathrm{Prob}^{|\Psi>}$ $(Q_2 = q^\alpha / Q'_1 = q^\beta)$ where $Q'_1 = Q' \otimes I$, $Q'_2 = I \otimes Q'$ and Q' is a self-adjoint operator *distinct* from Q. This generalization also applies to the other violations of PII discussed in this section.

[11] See D'Espagnat [1976] pp. 58–61 for a discussion of mixed states arising in this way. He calls them 'improper' mixtures.

[12] A similar point is made in van Fraassen [1984]. See also Margenau [1944] and [1950].

[13] It should be stressed that these relational properties expressed by the conditional probabilities in no way supervene on the monadic properties expressed by the marginal distributions. In the terminology of Teller [1986] they are *inherent* relations.

(2) There is another sort of relational property we might consider, expressed by comparing

$$\text{Prob}\ (Q_2 = q^{\alpha}/Q_2 = q^{\beta}) = \delta_{\alpha\beta} \tag{13}$$

with Prob $(Q_1 = q^{\alpha}/Q_2 = q^{\beta})$ given by (12).

These relational properties of particle 2 to itself as compared with relations of particle 1 to particle 2 we reject as vindication of PII, for the same argument as we discussed in section 1 for ruling out illegitimate trivializations of PII. The purported vindication of PII again depends on regarding (13) as a *monadic* property of particle 2 whereas it is in fact a relational property of particle 2 to itself, which is also true as a relation of particle 1 to *itself.*

(3) If we write $\alpha = \beta$ in equation (12) then we indeed find for fermions the result

$$\text{Prob}^{|\Psi>}\ (Q_1 = q^{\alpha}/Q_2 = q^{\alpha}) = 0$$

This shows that if actualization of Q_2 gives a certain value then there is zero probability that a concurrent actualization of Q_1 will yield the same value. This is the real significance of the Exclusion Principle, but has no bearing on PII, if we adhere to the orthodox view that actualizations do not correspond to antecedently existing possessed values.

(4) This brings us to our final comment. In hidden-variable reconstructions of QM, the circumstance demonstrated in point 3 above, *would* lead us to regard PII as vindicated for fermions (assuming that actualization results merely reveal preexisting values).

We now turn to the case of bosons. It is often assumed that purported violation of PII depends on consideration of states such as (5) or (6) where both particles can indeed be attributed the same pure state[14]. Denoting the state (5) by $|\Phi>$, for example, we can easily obtain the following results corresponding to (11) and (12):

$$\text{Prob}^{|\Phi>}\ (Q_2 = q^{\alpha}) = |<q^{\alpha}|a^{r}>|^2 \tag{11$'$}$$

and

$$\text{Prob}^{|\Phi>}\ (Q_1 = q^{\alpha}/Q_2 = q^{\beta}) = \text{Prob}^{|\Phi>}\ (Q_2 = q^{\alpha}/Q_1 = q^{\beta}) = |<q^{d}|a^{r}>|^2 \tag{12$'$}$$

So, as we might expect, both monadic and relational properties are the same for the two particles. But it should be noted that this conclusion is also true for the state (7), where two different states are involved. In this case the results (11) and (12) apply with the minus sign in front of the 'interference' term in (12) replaced by a plus sign.

Finally, we make a brief comment on the case of paraparticles. Here there *do* exist states for which the monadic properties of all the separate particles are not the same, but equally there are possible paraparticle states for which PII *is* violated in the same way as for bosons and fermions.

---

[14] See, for example, the discussion of PII for bosons given by Cortes [1976], Barnette [1978], Ginsberg [1981] and Teller [1983].

As an example consider the following normalized state for three parabosons of order 2[15]:

$$|\Psi'> = \frac{1}{\sqrt{2}}(|a^r> \otimes |a^r> \otimes |a^s> - |a^s> \otimes |a^r> \otimes |a^r>) \quad (14)$$

where $|a^r>$ and $|a^s>$ are two distinct one-particle states and triple tensor products are written in the sequence of particle labels 1, 2 and 3.

Denoting $Q \otimes I \otimes I$ by Q_1, $I \otimes Q \otimes I$ by Q_2 and $I \otimes I \otimes Q$ by Q_3, we obtain for the triple joint distribution

$$\mathrm{Prob}^{|\Psi'>} (Q_1 = q^\alpha \& Q_2 = q^\beta \& Q_3 = q^\gamma)$$

$$= |<q^\alpha| \otimes <q^\beta| \otimes <q^\gamma|(|\Psi'>)|^2$$

$$= \tfrac{1}{2}[|<q^\alpha|a^r>|^2 \cdot |<q^\beta|a^r>|^2 \cdot |<q^\gamma|a^s>|^2$$

$$+ |<q^\alpha|a^s>|^2 \cdot |<q^\beta|a^r>^{2} \cdot |<q^\gamma|a^r>|^2$$

$$-2\mathrm{Re} <a^r|q^\alpha> <q^\alpha|a^s> <a^r|q^\beta> <q^\beta|a^r> <a^s|q^\gamma> <q^\gamma|a^r>] \quad (15)$$

From (15) we find immediately the marginal distributions

$$\mathrm{Prob}^{|\Psi'>} (Q_1 = q^\alpha) = \mathrm{Prob}\ |\Psi'> (Q_3 = q^\alpha)$$

$$= \tfrac{1}{2}(|<q^\alpha|a^r>|^2 + |<q^\alpha|a^s>|^2) \quad (16)$$

while

$$\mathrm{Prob}^{|\Psi'>} (Q_2 = q^\alpha) = |<q^\alpha|a^r>|^2 \quad (17)$$

Thus particles 1 and 3 have the same monadic properties expressed by the marginal distributions but these differ from the monadic properties of particle 2.

Let us now show that particles 1 and 3 also have the same relational properties with respect to both the remaining particles.

We easily find that

$$\mathrm{Prob}^{|\Psi'>} (Q_1 = q^\alpha/Q_3 = q^\gamma) = \mathrm{Prob}^{|\Psi'>} (Q_3 = q^\alpha/Q_1 = q^\gamma)$$

$$= [|<q^\alpha|a^r>|^2 \cdot |<q^\gamma|a^s>|^2 + |<q^\alpha|a^s>|^2 \cdot |<q^\gamma|a^r>|^2$$

$$-2\mathrm{Re} <a^r|q^\alpha> <q^\alpha|a^s> <a^s|q^\gamma> <q^\gamma|^r>]/$$

$$[|<q^\alpha|a^r|^2 + |<q^\alpha|a^s>|^2] \quad (18)$$

Furthermore

$$\mathrm{Prob}^{|\Psi'>} (Q_1 = q^\alpha/Q_2 = q^\beta)$$

---

[15] Compare Hartle and Taylor [1969] for details of how to construct paraparticle states. The state (14) is obtained from their (2.5) by identifying two of the one-particle states.

$$= \mathrm{Prob}^{|\Psi'>} \ (Q_3 = q^\alpha / Q_2 = q^\beta)$$

$$= \tfrac{1}{2}[| < q^\alpha | a^r > |^2 + | < q^\alpha | a^s > |^2] \qquad (19)$$

and finally

$$\mathrm{Prob}^{|\Psi'>} \ (Q_1 = q^\alpha / Q_2 = q^\beta \ \& \ Q_3 = q^\gamma)$$

$$\mathrm{Prob}^{|\Psi'>} \ (Q_3 = q^\alpha / Q_2 = q^\beta \ \& \ Q_1 = q^\gamma)$$

$$= [| < q^\alpha | a^r > |^2 \cdot | < q^\beta | a^s > |^2 \cdot | < q^\gamma | a^r > |^2$$

$$+ | < q^\alpha | a^s > |^2 \cdot | < q^\beta | a^r > |^2 \cdot | < q^\gamma | a^r > |^2$$

$$-2\mathrm{Re} < a^r | q^\alpha > \ < q^\alpha | a^s > \ < a^s | q^\beta > \ < q^\beta | a^r > \ < a^r | q^\gamma > \ < q^\gamma | a^r >]/$$

$$| < q^\beta | a^r > |^2 [| < q^\gamma | a^r > |^2 + | < q^\gamma | a^s > |^2] \qquad (20)$$

These results show that PII is violated for particles 1 and 3 in the state $|\Psi' >$, even in its weakest form, that includes all the relevant relational properties.

5 CONCLUSION

There are two main conclusions of this paper. Firstly that indistinguishable particles in QM can be treated as individuals, but secondly, if they are so treated, then, on the most plausible reading of what constitutes a property of a quantal particle, even the weakest form of PII, including both monadic and relational properties, is violated both for bosons and fermions, and indeed for higher-order paraparticles.

It should be noted that if quantal particles *are* individuals, then their individuality must be conferred by TI. STC is not in general available in QM, since particles do not move in well-defined trajectories, so the question of spatio-temporal continuity of trajectory does not arise. The only exception to this is where the one-particle states involve well-defined wave packets, which diffuse sufficiently slowly over time, as would be possible for the classical limit of sufficiently massive particles.

But it is clear that in the case of macroscopic bodies, where STC can be used to label the bodies, the STC criterion actually conflicts with the TI individuation of the elementary particles composing the body. To be strict every electron, for example, partakes of the state of every other electron in the universe, according to the antisymmetrization requirement!

But notice, that under conditions where the 'interference' term in (10) can be neglected, then the state $|\Psi >$ behaves like a *proper* mixture of states in which particle 1 is in state $|a^r >$ and particle 2 in state $|a^s >$ and the permuted state in which particle 1 is in state $|a^s >$ and particle 2 in state $|a^r >$, with equiprobable weights for the two component states in the mixture. So under these conditions the state (8) behaves like an equiprobable mixture of the states (3) and (4). In other words, when 'interference' can be neglected, we recover the same possibilities for states as in classical physics, where states (1) and (2) would anyway be eliminated by IA assuming them to be maximally specific.

But, of course, ontologically speaking, 'interference' is never strictly absent. That, after all, is what constitutes the 'problem of measurement' in QM, so the involvement of every electron with the state of every other electron in the universe, although negligible for practical purposes, remains an ontological commitment of QM, under the interpretation where the particles are treated as individuals. If this sounds too bizarre to be acceptable, it provides another argument for preferring the treatment of indistinguishable particles along the lines provided by quantum field theory.

In this paper we have been concerned with conceptual possibilities, rather than what is most reasonable to believe about the ontological status of elementary particles.

Department of Philosophy, and Centre of Logic,
Epistemology and History of Science, Universidade Estadual de Campinas
Campinas, S.P. Brasil

History and Philosophy of Science,
King's College London (KQC) England

REFERENCES

Barnette, R. L. [1978]: 'Does Quantum Mechanics Disprove the Principle of the Identity of Indiscernibles?', *Philosophy of Science*, 45, pp. 466–70.

Cortes, A. [1976]: 'Leibniz's Principle of the Identity of Indiscernibles: A False Principle', *Philosophy of Science*, 43, pp. 491–505.

D'Espagnat, B. [1976]: *Conceptual Foundations of Quantum Mechanics*, 2nd Ed. Benjamin.

Ginsberg, A. [1981]: 'Quantum Theory and the Identity of Indiscernibles Revisited', *Philosophy of Science*, 48, pp. 487–91.

Hartle, J. B. and Taylor, J. R. [1969]: 'Quantum Mechanics of Paraparticles', *Physical Review*, 178, pp. 2043–51.

Hoy, R. C. [1984]: 'Inquiry, Intrinsic Properties and the Identity of Indiscernibles', *Synthese*, 61, pp. 275–97.

Ishiguro, H. [1976]: 'Leibniz's Theory of the Ideality of Relations', in H. G. Frankfurt (ed.): *Leibniz: A Collection of Critical Essays*, Notre Dame University Press, pp. 191–213.

Leibniz, G. W. [1981]: New Essays on Human Understanding. Translated and edited by P. Remnant and J. Bennett, Cambridge University Press.

Lucas, J. [1984]: *Space, Time and Causality*. Oxford University Press.

Margenau, H. [1944]: 'The Exclusion Principle and its Philosophical Importance', *Philosophy of Science*, 11, pp. 187–208.

Margenau, H. [1950]: *The Nature of Physical Reality*. McGraw-Hill.

Messiah, A. M. L. and Greenberg, O. W. [1964]: 'Symmetrization Postulate and its Experimental Foundation', *Physical Review*, 136B, pp. 248–67.

POST, H. R. [1963]: 'Individuality and Physics', *The Listener*, 10th October. Reprinted in *Vedanta for East and West*, 132, (1973), pp. 14–22.

REDHEAD, M. L. G. [1983]: 'Quantum Field Theory for Philosophers', in P. D. Asquith and T. Nickles (eds.): *Proceedings of the 1982 Biennial Meeting of the Philosophy of Science Association, Vol. 2*, pp. 57–99.

REDHEAD, M. L. G. [1988]: 'A Philosopher Looks at Quantum Field Theory', in H. R. Brown and R. Harré (eds.): *Philosophical Conceptual Foundations of Quantum Field Theory*. Oxford University Press. pp. 9–23.

SHADMI, Y. [1978]: 'Teaching the Exclusion Principle with Philosophical Flavour', *American Journal of Physics* 46, pp. 844–8.

TELLER, P. [1983]: 'Quantum Physics, the Identy of Indiscernibles and Some Unanswered Questions', *Philosophy of Science* 50, pp. 309–19.

TELLER, P. [1986]: 'Relational Holism and Quantum Mechanics', *British Journal for the Philosophy of Science* 37, pp. 71–81.

VAN FRAASSEN, B. [1984]. 'The Problem of Indistinguishable Particles', in J. T. Cushing, C. F. Delaney and G. M. Gutting (eds.): *Science and Reality: Recent Work in the Philosophy of Science*. University of Notre Dame Press, pp. 153–72.

Brit. J. Phil. Sci. 39 (1988), 247–261 *Printed in Great Britain*

Zeeman–Göbel Topologies

ADRIAN HEATHCOTE

ABSTRACT

In 1967 E. C. Zeeman argued that the Euclidean (i.e. manifold) topology of Minkowski space-time should be replaced by a strictly finer topology that was to have a closer connection with the indefinite metric. This proposal was extended in 1976 by Rudiger Göbel and Hawking, King and McCarthy to the space-times of General Relativity. It is the purpose of this paper to argue that these suggestions for replacement misrepresent the significance of the manifold topology and overstate the necessity for a finer topology. The motivation behind such arguments is a realist view of space-time topology as against (what can be construed to be) the instrumentalist position underlying some of the suggestions for replacement.

The currently dominant view of space-time geometry—a view that lies behind most realist, non-conventionalist discussions of the subject—is that at each structural level of the space-time, properties are well-defined, whether derived immediately from basic physical postulates or entailed by other structural properties. The following list (H) is a common way of dividing the space into structural levels and is taken over from classic treatments of differential geometry.

(H) (i) Metric tensor field
(ii) affine connection
(iii) conformal structure
(iv) differentiable manifold
(v) topology

Occasionally, as in the discussions derived from the work of Ehlers, Pirani and Schild [1972] (also Ehlers and Schild [1972]), a projective structure may be included, and in specialized circumstances a spinor structure above (i). Similiarly, we might have included the curvature structure, either in the form of the Riemann tensor or the sectional curvature function. Whatever fine adjustments are made, however, (H) represents a convenient breakdown of the structural properties of space-times. Levels (i) to (iii) represent geometrical objects and (iv) and (v) are the structures of the manifold itself. The method that results from considering tensor bundles, connections on bundles and prolongations is not essentially different in this context from that which results in (H) in more conventional treatments.

The rationale for (H) cannot be that each level below (i) has a transformation group that is a subgroup of the transformation group of that level which precedes it, since for most spaces encountered in the General Theory of Relativity (henceforth GTR)—and certainly those that arise under realistic circumstances—the group of isometries will be equivalent to the group of affine transformations or even the group of conformal isometries. [The theorems that connect the isometry group with the group of affine transformations on positive-definite Riemannian manifolds are well-known (see Kobayashi and Nomizu [1963] Chapter VI) but the corresponding theorems for Lorentz-signature manifolds are not completely known. There are however some suggestive theorems in De Frise-Carter [1975] that are sufficient to justify the above remark].

A better explanation of (H) is in terms of uni-directional determination of a structure lower on the list by one that is higher. On this view the affine connection is weaker than the metric because fixing the metric will determine a unique affine connection (the Levi-Civita 'metric connection') but the converse does not hold. The same holds for the conformal structure and the metric. Similarly, the manifold can be defined without defining a metric but the metric cannot be defined without the existence of a manifold structure. The manifold carries with it a natural topology—the manifold topology—but this topology does not fix a smooth differentiable structure. On this view then the hierarchy (H) is justified by this uni-directional determination running down through the structural levels and taking as its base the manifold topology as the deepest non-trivial structure. (I will not elaborate—since it is besides present purposes—on the necessary caveat that it is not quite true that (H)ii determines (H)iii. That aside, (H) stands).

A justification for this interpretation of (H), one that may be broadly characterized as realist in spirit, is that if each level determines deeper levels, then, given that the metric tensor is uniquely defined by physical processess, the structures which are determined by the metric tensor must be similarly unique and well-defined. For nomenclature's sake I'll call this position *trickle-down realism*. Such a view, it may be argued, has been implicit in much of the post-conventionalist analyses of GTR. In particular this interpretation of (H) has been committed to the existence of a unique topology for space-time which—quantum speculations aside—has been taken to be the manifold topology. This view of the manifold topology has been so much a part of an orthodoxy that it is rare (except in recent years) to have the manifold topology referred to as anything other than 'the topology'.

This realist view of (H) at all structural levels has suggested a *constructive unity* to space-time, a position which has been supported by the work of Ehlers, Pirani and Schild who have attempted to find a precise physical motivation for accepting each of the geometrical structures of the space-time. The fact that these structures have, for a long time, been considered to be both well-defined and unique has led to the increasing acceptance of realism (with respect to space-time) over instrumentalism and conventionalism. If, for example, a space had a well-defined affine connection but it was impossible to precisely fix a manifold structure (or, more precisely, where one could be abstractly fixed, say from the existence of convex neighbourhoods, but which was contradicted directly by a physical argument) then realism with respect to that geometry may no longer be considered the most acceptable position. Similarly, if a manifold could be uniquely defined but a metric tensor could not, then this may well be taken as an indication that the function of space-time in the physical theory had altered to the point where we ought to question whether even the manifold structure was strictly required. (This situation bears an intentional similarity to Adolph Grünbaum's thesis of Metrical Amorphousness. Grünbaum, of course, did not construe the implications of this position as I have suggested. To explain this we accept Hilary Putnam's argument that Grünbaum did not do so because he had an essentialist position about topology: he regarded it as all a (continuous) space really was.) Well-defined, unique properties have long been taken to be an important component of realism—as is evident from the recurrent threats to realist positions arising from underdetermination arguments. If we have two competing attributions of properties to an entity that we are, in principle, unable to resolve then we are inclined (whether motivated by methodological principles or empiricist prejudice) to make the ontological inference that there is no precise property to capture.

These two aspects of realism, namely uniqueness and well-definedness, have strong arguments supporting them with global topological properties, e.g., connectibility, compactness and orientability. If there is a certain amount of matter-energy in the universe distributed in a certain way then the Einstein field equations should be solvable to give a precise metric and a precise global topology. To be conventionalist about the latter would require being conventionalist about the former. Of course if there is no accurate estimate for the former then there may indeed be no accurate model for the latter but cosmology is at least predicted on the connection between the two. We have good reason to believe therefore in well-defined and unique global properties. This paper is concerned rather with challenges to the uniqueness of local properties of the topology.

In 1967 E. C. Zeeman discussed the possibility of defining a new topology for Minkowski space-time [1967]. In that paper Zeeman stated that,

> It is customary to think of M (Minkowski space) as having the topology of real 4-dimensional Euclidean space, although there are reasons why this is wrong. (p. 161)

Zeeman criticised the usual topology on two grounds.

> (1) The 4-dimensional Euclidean topology is locally homogeneous, whereas *M* is not; every point has associated with it a light cone separating space vectors from time vectors. (2) The group of all homeomorphisms of 4-dimensional Euclidean space is vast, and of no physical significance. (p. 161)

Zeeman proposed a new topology which, in contrast to the Euclidean topology, had what were taken to be the desirable properties that it was not locally homogeneous since it encoded the light cone structure, and the group of homeomorphisms of the new topology was generated by the inhomogeneous Lorentz group plus dilatations. The new topology also has the supposedly more physical consequence that the track of a light ray is not a continuous path but is discrete. Zeeman remarks:

> . . . the induced topology on a light cone is subtle, and a function on Minkowski Space that is not continuous in the 4-dimensional Euclidean topology because of discontinuity on the light cone, might well become continuous in the fine topology. (p. 162)

A similar proposal to that of Zeeman's was made simultaneously by the Russian mathematician Cel'nik [1968]. The fine topologies were then generalized in 1976 by Rudiger Göbel to the space-times of GTR [1976a], [1976b] and an additional topology was proposed for strongly causal space-times by Hawking et al [1976] in the same year. I will refer to all of these topologies collectively as *Zeeman-Göbel topologies* even though they may differ from one another in important ways.

We now give some background detail on the notions of topology that will be important in the ensuing discussion and sketch the basic concepts of the Zeeman-Göbel topologies.

If X is a set and T is a family of subsets of X then T is a topology on X iff

(i) both $\emptyset \, \varepsilon \, T$ and $X \, \varepsilon \, T$,

(ii) if $\{T_i \mid i \, \varepsilon \, \lambda\}$ is a class of elements of T then $\left(\bigcup_{i \varepsilon \lambda} T_i\right) \varepsilon T$,

(iii) if $\{T_i \mid i \, \varepsilon \, \Gamma\}$ is a finite subclass of elements of T then $\left(\bigcap_{i \varepsilon \Gamma} T_i\right) \varepsilon \, T$.

The pair (X,T) is then termed a *topological space*. Informally, a topology on a set X is a subset of the power set of X containing the empty set and X itself which is closed under arbitrary unions and finite intersections.

A topological space *M* is an n-dimensional *manifold* if *M* is locally homeomorphic to R^n, the n-times Cartesian product of R (the real numbers) with itself. Specifically, if *U* is an open set of *M* and ϕ: $U \rightarrow R^n$ is a homeomorphic map that takes *U* onto an open subset of R^n then ϕ is called a coordinate map and (U, ϕ) is a *chart* (or a coordinate system). The dimension of M is then the integer n. Let $\{(U_i, \phi_i) \mid i \varepsilon \Theta\}$ be a set of charts on M such that the set $\{U_i \mid i \varepsilon \Theta\}$ forms an open cover on M.

A map f from an open set $O \subset R^n$ such that $f: O \rightarrow R^m$ is said to be *differentiable of class C^r* (or simply a C^r map) if all partial derivatives up to order r of f exist and are continuous. A C^0 map is simply a continuous map and a C^∞ map is C^r for all values of r. Two charts (U_i, ϕ_i) and (U_j, ϕ_j) with $U_i \bigcap U_j \neq \emptyset$ have a C^r overlap if, for $\phi_i (U_i \bigcap U_j)$ and $\phi_j (U_i \bigcap U_j)$ open in R^n,

$$\phi_j \circ \phi_i^{-1}: \phi_i (U_i \bigcap U_j) \rightarrow \phi_j (U_i \bigcap U_j)$$

is differentiable of class C^r and $\phi_i \circ \phi_j^{-1}$ is likewise. A C^r atlas on M is one in which all charts in the set $\{(U_i, \phi_i) \mid i \varepsilon \Theta\}$ with non-empty intersection have a C^r overlap. A C^r atlas is said to be *complete* or *maximal* if every chart which can have a C^r overlap with every other chart in the atlas is also contained within the atlas. A complete C^r atlas Φ is then said to be a *C^r differential structure* on M and the pair (M, Φ) is a *C^r differentiable manifold.*

The manifold can be considered to be an ordered triple (X, T, Φ) consisting of an underlying point set X, a second countable, usually Hausdorff and paracompact, topology T on the set X and a complete C^r atlas Φ. A C^0 manifold is often called a *topological manifold.* (More information on topology, including the definitions of the above properties, can be found in Hu [1964] and on manifolds in Bishop and Goldberg [1968].)

Given a manifold $M = (X, T, \Phi)$ it is obvious that we can 'forget' the C^r differentiable structure in the following sense: given a C^k differentiable structure on a manifold then for every r such that $0 < r < k$ there is a compatible C^r atlas on M - simply because any C^k coordinate change is automatically a C^r coordinate change if $r < k$. However the effect of doing this is to enlarge the maximal atlas to include all those charts that are C^r related. Thus every C^r manifold is automatically a C^0 manifold as a result of 'forgetting' the extra degrees of differentiability in the coordinate changes (C^r overlap). We thereby recover a unique topology from the manifold, namely the second element in the triple (X, T, Φ). This much seems trivial—we forget the smoothness but retain the continuity: we have done nothing more than stress the compatibility between the structures. There is, however, a substantive point here as well. Rather than define a differential structure *on a topological space* we could just as well have defined it on the set X (though the definition would have required more manipulation). The atlas Φ would be defined by the fact (as before) that coordinate changes are C^r and there will be a unique topology on X that makes the atlas Φ C^r. It thus makes little difference from the point of view of definition whether one assumes a topological space on which the differential structure is defined or whether the differential structure is defined on a set and the topology taken as an induced structure. In either case the topology is unique to the manifold and it is this topology that will be referred to hereafter as the *manifold topology*.[1] Henceforth the manifold will be represented by the symbol '*M*' alone.

There are other important ways in which a topology may be induced on the set of points underlying a space. One such way is the topology that is induced by a distance function on a space.

We first note that a *local basis* or a *neighbourhood basis* for a topological space (X,T) is a collection of basis neighbourhoods for each point $x \in X$. A basis neighbourhood for a point $x \in X$ is a collection N_x^B of neighbourhoods of x such that every neighbourhood of x contains one of the basis neighbourhoods in N_x^B. And, finally, a neighbourhood of $x \in X$ is any subset $N \subset X$ such that $x \in \text{Int}(N)$—the largest open subset of *N*.

Given the distance function $d: X \times X \rightarrow R$ satisfying the usual axioms (positivity, symmetry, nondegeneracy and the triangle inequality) the metric topology induced by d on X can be defined in terms of a basis of neighbourhoods consisting of open balls (where an open ball centered at the point x and radius $\varepsilon > 0$, denoted by $N_\varepsilon(x)$, is the set $\{y \mid d(x,y) < \varepsilon\}$). It is an important point that the n-dimensional Euclidean metric

$$d(x,y) = [\sum_{i=1}^{n} (x_i - y_i)^2]^{\frac{1}{2}}$$

induces a topology on the underlying set[2], the Euclidean topology, which is *equivalent* to the manifold topology. Indeed, since a positive-definite Riemannian metric can be used to define a unique distance function, the metric topology of every Riemannian manifold is equivalent to the underlying manifold topology of the space. It is crucial to the motivating of the Zeeman-Göbel topologies that this equivalence (manifold topology = metric topology) does not hold for the case of Lorentz signature metrics on a given manifold.

We now define Zeeman's fine topology on Minkowski space-time. Let the pair (M, η), where $\eta = (dt)^2 - (dx_1)^2 - (dx_2)^2 - (dx_3)^2$ is the metric expressed in the usual cartesian coordinates on R^4, denote Minkowski space-time and $Q = t^2 - x_1{}^2 - x_2{}^2 - x_3{}^2$ be the characteristic quadratic form on (M, η) associated with the metric η. We denote straight (Euclidean) timelike and null lines by the symbols R_T and R_L respectively, and we denote spacelike hyperplanes by $R_s{}^3$. Zeeman defines the following three cones at every point: the Light Cone $C^L(x) = \{y \mid Q(y-x) = 0\}$, the Time Cone $C^T(x) = \{y \mid y = x \text{ or } Q(y-x) > 0\}$ and the Space Cone $C^s(x) = \{y \mid y = x \text{ or } Q(y-x) < 0\}$.

---

[1] Since 1983 the understanding of the differential topology of four-dimensional manifolds has been considerably advanced (and in some respects thrown into turmoil) by S. K. Donaldson's work on the topology of the space of solutions of the self-dual Yang-Mills equations on an *SU*(2) bundle over a Riemannian manifold [1983]. One startling consequence of Donaldson's main theorem is that, when combined with recent work by Michael Hartley Freedman [1982] it can be shown that R^4 possesses exotic, i.e. non-standard, differentiable structures: something that is known not to be true for any other R^n. Some of these exotic differentiable structures have been explicitly constructed (Gompf [1983]). Penrose (in [1984]) has suggested that Donaldson's theorem may provide a new model of Quantum Gravity, (c.f. also Atiyah [1983]).

[2] Many distinct, non-trivially different, metrics can induce the same topology on an underlying set. For example the usual Euclidean metric of R^2 and the taxi cab metric given by $d(x,y) = |x_1 - x_2| + |y_1 - y_2|$ both induce the same topology (c.f. Voxman & Christensen [1977] p. 20—the reader should be warned that there is a notational snarl in their proof and the descriptions of the metrics). Similarly, if (X,d) is a metric space the bounded metric $d' = \frac{d}{(1+d)}$ will define the same topology on X.

The group G is the group of automorphisms of (M, η) consisting of the Lorentz group + translations + dilatations and G_0 is the subgroup of G that preserves the time ordering of the timelike vectors. Given a specific timelike line R_T and spacelike plane $R_s{}^3$ the entire class of such lines and planes can be generated by elements of the group G.

The fine topology F on (M, η) is the finest topology (maximal number of open sets) such that the induced topology on all R_T and $R_s{}^3$ is the same as that induced by the Euclidean topology on M. Fine neighbourhoods $N_\epsilon^F(x)$ can be defined in terms of their Euclidean counterparts $N_\epsilon^E(x)$ taking the union of the space and time cones at x and forming the intersection between the resulting set and $N_\epsilon^E(x)$. The N_ϵ^F do not form a basis of neighbourhoods in the fine topology. This fine topology on Minkowski space-time is not paracompact or normal though it *is* Hausdorff and connected. Since the null cone points are not in the fine neighbourhood $N_\epsilon^F(x)$—with the exception of x itself, of course—the path of a photon is discrete in the fine topology.

Since the fine topology has more open sets than the manifold topology there are disjoint sequences $S = \{p_n\}$ such that the sequence $\{p_n\}$ converges in the manifold topology but not in the fine topology, i.e. $p_n \xrightarrow{m} P$ but $P_n \not\xrightarrow{F} P$. Zeeman terms these 'Zeno sequences'. Zeno sequences are used to prove that a continuous map of the unit interval into (M, η) endowed with the fine topology has an image which is piecewise linear, thus imitating a free particle undergoing collisions.

The Zeeman topologies defined on GTR space-times by Rudiger Göbel are similar to, though necessarily more sophisticated than, the fine topology defined above. The family of topologies described in Göbel [1976a] and denoted generically as $\mathfrak{Z}$, retain the idea of continuity on space-like hypersurfaces and (if there is no electro-magnetic field present), $\mathfrak{Z}$-continuous possible world lines are future directed, timelike, piecewise geodesic curves. Göbel's significant advance over Zeeman is his incorporation of external 'physical' fields into the systems that are to determine the $\mathfrak{Z}$ topologies. (Göbel only mentions the electro-magnetic field but there is no reason to think that the same could not be done for the strong and weak fields with a resultant coding of the quantum structures into a modified space-time structure—but though this may be mathematically interesting it is the purpose of this paper to argue the view that it may not mark an ontologically significant strategy.)

In his [1976b] Göbel responded to criticism from Hawking, King and

McCarthy [1976—henceforth HKM] and defined a new set of fine topologies (one of which was the path topology *P* of HKM) which broadened the class of $\mathfrak{Z}$-continuous world lines and dropped the condition that spacelike hypersurfaces are continuous in the fine $\mathfrak{Z}$ topology.

The path topology *P* defined by HKM is the finest topology such that the induced topology on all continuous timelike curves is the same as that induced by the manifold topology on those curves. (Here and throughout I won't scruple about the distinction between paths and curves since it's of no consequence for my argument—the mathematical literature seems often similarly ambivalent about the distinction.)

A se P $\subset M$ is open in the *P*-topology iff for every timelike curve there is a set *O*, open in the manifold topology, such that $P \cap t = O \cap t$. The path topologies defined in Göbel [1976b] do not differ significantly from the path topology *P* given by the condition above. The topology that Göbel favours restricts the admissible timelike curves t to those that have a reasonable differentiability, say C^r, where the manifold is also taken to be C^r. However this restriction, physically motivated though it is, seems unnecessarily strong. It seems reasonable therefore to concentrate on *P* and introduce restrictions on timelike curves only if the discussion warrants it. The path topology *P* is not paracompact, though unlike $\mathfrak{Z}$, it *is* first countable. *P* is not locally compact, normal, or regular and, by a theorem of A. H. Stone [1948], since *P* is not paracompact, it is not metrizable either; similarly, since *P* is T_1 and unmetrizable, it is not pseudo-metrizable[3], c.f. Kelley [1955]. *P* is Hausdorff, seperable, connected, locally connected, path connected and locally path connected.

Both the *P* and $\mathfrak{Z}$ type topologies are regarded as having various advantages over the manifold topology—most notably that automorphisms of the fine topologies are respectively conformal and homothetic transformations of the space-time on which they are defined. It is not, however, the technical results concerning the fine topologies that are in question or that are predominately at issue in this paper. Rather, the focus of attention is on the *nature of topology itself*, in its physical manifestations, in the light of these suggested alternative topologies for space-times.

It would appear from his introductory remarks, quoted above, that Zeeman regards the manifold topology of Minkowski space-time as not the correct topology and that the fine topology is an appropriate *replacement* for the Euclidean topology. It is unclear, though, in what sense the manifold topology *can* be replaced. It is certainly true that the open sets of the fine topology, since they are defined on the basis of timelike and spacelike lines, encode information which is present in the upper levels of structure, but still, the manifold topology does not seem to be redundant in the way Zeeman suggests. Göbel makes similar motivating remarks to those of Zeeman, commenting that,

---

[3] A pseudo-metric in this sense obeys the usual axioms for a distance function (see Kelley [1955] p. 118) with the exception that $d(x,y) = 0$ does not require $x = y$. It should be noted that on this definition a Lorentz metric is *not* a pseudo-metric.

> the ordinary Euclidean topology $\mathfrak{T}$ . . . is artificial from both a mathematical as well as the physical point of view. The topology $\mathfrak{T}$ (on a Minkowski space) can be understood primarily from its history: People were mostly concerned with Riemannian spaces (with a positive metric) for which $\mathfrak{T}$ is a very natural topology, and not with spaces provided with a pseudo-Euclidean metric or in particular with a Lorentz metric. [1976a] p. 290

Zeeman's and Göbel's remarks give the impression that there is a *unique correct* topology for spaces equipped with a Lorentz metric and that this is one of the fine topologies outlined above. But even if the fine topologies are physically well motivated the manifold topology must still exist as a definite structure of the space-time since the manifold itself is just such a definite structure and still exists. The manifold topology is present as long as the manifold is present. What options does the *replacement thesis*—the view apparently advocated by Zeeman and Göbel—allow us? I think that there are two possibilities, both untenable.

(1) Both the manifold structure and the accompanying manifold topology are to be replaced by one of the fine topologies as the sub-structure of the space-time.
(2) The manifold structure is to be retained but the manifold topology is to be replaced by one of the fine topologies.

In both (1) and (2) the choice of which fine topology is to stand as replacement depends entirely on which is to be considered as the most physically well-motivated.

I think (1) is untenable because the linear connection and the metric tensor field both depend to such a great degree on the existence of a manifold structure, so much so that neither could exist without one. And if there is no metric tensor (or metric connection) it is not possible to define the geodesic curves (since curves are geodesics only with respect to a specification of a metric or a connection) and hence it is not possible to define the fine topologies. Furthermore, the the tensor bundles and connections have certain smoothness requirements that depend upon the smoothness of the manifold structure: the manifold is highly indispensible, at least so long as the geometrical objects are retained in their present form. But using them in their present form is what the fine topologies are about. Therefore the fine topologies cannot replace the manifold structure. This leads on to (2).

In describing the differential structure of manifolds I attempted to show that they do not bear an accidental relationship to their own topologies. A differentiable function is of necessity a continuous function. A differentiable manifold is of necessity a topological space. Indeed, a differentiable manifold is little more than a topological space that is locally homeomorphic to R^n, for some n. To say, therefore, that the manifold topology can be replaced by a fine topology that leaves the differential structure in place makes no sense. The differential structure is a refined characterization of the topological space not something apart from it. It does not seem possible then that (2) could be true. The manifold topology cannot be replaced by a fine topology in such a way as to leave the manifold in place.

A supporter of the Zeeman-Göbel topologies may feel that the above argument does not do justice to the motivation for alternative topologies and that something like the replacement thesis may be true if we relax the constraints on what it is for a topology to obtain. In sympathy with this view I will put the strongest case that I can think of for what I will call the *weakened replacement thesis* (WRT).

An advocate of WRT might argue as follows: the above argument against the replacement thesis assumes an either/or position on topology-1) either the manifold topology exists or a Zeeman-Göbel topology exists, but not both, 2) the manifold topology must exist, since the manifold itself must, therefore none of the Zeeman-Göbel topologies can be the real topology of a space-time. This either/or position is untenable. Perhaps the real sense in which a topology can be said to exist is more a matter of it being definable. This is something that does not come up in the case of (positive-definite) Riemannian manifolds—not, at least, for practical purposes though, of course, alternative topologies are still definable—because in that case the metric topology is the same as the manifold topology. Since this equivalence does not hold in pseudo-Riemannian manifolds we must adopt a new attitude to the existence of a topology. We must look at all definable topologies on a space-time and choose one that is parsimonious and best reflects the physical structure of the space-time. After all, *you* say (in the above argument against the replacement thesis) that the manifold topology must exist just because the manifold exists, surely then by parity of reasoning the Zeeman-Göbel topologies must exist just because the metric-causal structure exists. Why should trickle-down realism give one any reason to prefer the manifold topology over the fine topologies—if realism trickles down for one then it equally trickles down for the other . . .

So I imagine a defender of WRT arguing (and clearly my statement of the position is easily extracted from the argument). I know of no-one who has run such a line but then the issue has never fully surfaced in the literature. At any rate I take the above to be consistent with (most of) Zeeman's and Göbel's motivating remarks.

I now want to say something against the above defence of WRT.

Firstly, it should be said that there is much in WRT that seems true. The point at which I think it misrepresents the situation is not in the definability of alternative topologies but in the specific arguments that it mounts in favour of the Zeeman-Göbel topologies. Let us agree that pseudo-Riemannian manifolds present us with a genuinely new situation from that which exists for positive-definite Riemannian manifolds, and also that the conformal part of the metric tensor can be used to define several denumerable classes of topologies in the former case that are different and strictly finer than the manifold topology. These finer topologies are not metric topologies but they are defined on the basis of aspects of the space-time geometry that derive from the conformal part of the metric. Why, to turn Zeeman's question on it's head, should we *care* about these topologies? There is certain structural information present at the level of the metric tensor so why should one want to duplicate it again at the level of the topology. I conjecture that part of the reason for Zeeman's own preference for the finer topologies is an unwarranted view that metric topologies, in and of themselves, have a more straightforward operational meaning than the manifold topology simply in virtue of the empirical status of the metric: a Grünbaumian position in reverse. I want to argue two (closely related) points: 1) that Zeeman and Göbel misrepresent the significance of the manifold topology, and 2) that their arguments, and those of HKM, on behalf of the fine topologies are too weak for the conclusion they want to draw. The second point first.

Göbel, following Zeeman's lead, makes the following remark: '. . . there are no experiments known to justify a Euclidean topology along light like geodesic lines. On the other hand the (induced) topology along the track of a photon is Euclidean in consequence of the manifold topology $\mathfrak{T}$' [1976] p. 290. He then goes on to say 'The fact that we do not have any geometric information along a light ray, can be observed of (sic) the Zeeman topology: . . . *The topology induced by the Zeeman topology on a light cone is discrete* (which means there is no topological 'information' on light rays). p. 290. But if, as Göbel says, we have *no* geometric information along a light ray then *a fortiori* we have no reason to prefer *any* topology along its path. (Perhaps this is an attempt to trade on the fact that emitted photons have discrete energies, but if it is why should quantization be taken to say anything at all about position.) The discreteness of photon tracks in the fine topologies gives no reason at all to prefer those topologies over the manifold topology. It is at best entirely neutral.[4] The second claim made on behalf of the fine topologies by Zeeman, Göbel and Hawking et al, is that automorphisms of the fine topologies correspond to physically significant groups, for example the orthochronous Lorentz group plus dilatations, the homothecy groups or the conformal groups. These authors do not indicate, however, just why a correspondence between the group of homeomorphisms and a group related to the automorphisms of the metric structure is desirable or necessary. After all, if we have a physically significant structure or geometric object field then there will be of necessity a corresponding automorphism group for that structure—even if it is degenerate. And conversely, if there is an automorphism group then—by definition—there must be something that is being preserved under the transformations. But what more do we gain by requiring that the group of *homeomorphisms* be *equivalent* to, say, *the group of conformal isometries*. As far as I can see we gain nothing. We are already convinced that the conformal part of the metric is physically significant but there is no reason to believe that the group of homeomorphisms either *becomes* significant when it equals the group of conformal isometries or that it needs to *be* significant. After all, in a general relativistic space-time the metric tensor is of the greatest physical significance but the group of isometries of the space-time is of no interest or at least very little (the significant exception to this is, of course, Minkowski space-time but any highly symmetric space-time, for example the exact solutions, will do)—and in physically realistic situations will be degenerate in that the only member of the group will be the identity map. The importance of a geometric object field is just a distinct matter from the significance of its automorphism group.

---

[4] I'm not convinced by Göbel's claim that we have *no* information along a light ray. Surely there would be some modification to optics required if light rays had discrete topologies. Perhaps there would be some detectable differences in interference experiments. I leave this as a tentative suggestion.

A third argument for the new topologies is that they incorporate a strictly 'physical' set of curves: in the case of HKM, for example, the 'set of continuous paths is the set of Feynman paths' (p. 175) and the topology that Göbel favours in [1976b] incorporates the *analytic* Feynman paths. The idea here is to try to cut down the class of continuous curves so as to be left with only those that have a physical significance, i.e. only those that can be realized as the paths of (possibly accelerated) particles. Thus curves that are too 'wild' in the sense that they are undifferentiable at too many points (such as Penrose's bad trips that have geodesic segments accumulating at a point), are eliminated from the class of continuous curves—continuous, that is, with respect to the fine topologies.

There is no doubt that cutting down the number of continuous paths with respect to a certain topology may have computational advantages in a small number of circumstances but there is no reason to think that changing the induced topology on non-physical curves gives one a reason for believing a fine topology actually *true*. After all, it is precisely the non-physical curves that one has no information about at all. Advocates of the HKM path topology might argue, of course, that it is precisely the fact that the new topologies alter only the topology on curves that are *not* physical that makes them an acceptable alternative. My point, however, is that this provides no motivation *by itself* for the change and since the other motivations also fail, or so I've argued, this third argument for the fine topologies has no force.

But, to return to the second claim made on behalf of the Zeeman-Göbel topologies, is it true, as Zeeman and Göbel believe, that the group of all homeomorphisms of a space-time is of no physical significance? The group of all homeomorphisms should be equivalent to—or very nearly equivalent to—the group of diffeomorphisms of a space-time, and this latter group is often taken to encapsulate the principle of General Covariance (see for example Guillemin and Sternberg [1983] Chapter 3). If this position is correct then far from being of no significance the group of homeomorphisms is perhaps more significant than any other group on the space-time: it contains as a large subgroup the Covariance group.

This takes me to point (1), the issue of the significance of the manifold topology. The dilemma for Zeeman and Göbel can be put thus: either their assumption that the significance of a structure is a reflection of the significance of its automorphism group is correct, in which case the manifold topology is physically significant and possibly more so than the group of homethecies or conformal isometries, or it is incorrect, in which case the fine topologies are physically unmotivated. But in either case the weakened replacement thesis provides no argument for the Zeeman-Göbel topologies.

The suggested connection between the principle of General Covariance and the group of diffeo-/homeomorphisms is not however the only reason for thinking the manifold topology significant. A more straightforward argument for the manifold topology, and one that is entirely physical in character, is based on the equivalence between the Alexandrov topology and the manifold topology (modulo the condition of strong causality). The Alexandrov topology is defined by specifying the basis given by $I^{+}(x)\bigcap I^{-}(y)$—the intersection of chronological past and future sets—for all points x and y. Since the Alexandrov topology is clearly physically significant the manifold topology must be also. We recover the physical significance of the manifold topology via the significance of chronological past and future sets. Once again, therefore, WRT gives no reason for preferring the fine topologies over the manifold topology. Quite the reverse, it gives us a reason to prefer the manifold topology.

I now want to make a few miscellaneous points concerning WRT and the arguments for the Zeeman-Göbel topologies..

Firstly, it may be thought that because spacelike curves are discrete, i.e. completely disconnected (as in the HKM Path topology) or light tracks are discrete (as in the Zeeman topologies) that these topologies are more *parsimonious*, that they are cutting down on our ontic commitments. Clearly, however, we have a change in *properties* not a change in *ontology*. If anything the ontic commitments should increase since discrete topologies contain more open sets than Euclidean topologies (though in the final analysis I think it makes little sense to talk this way). So there is no obvious gain in parsimony.

Secondly, as I suggested above, if we have no information along light paths, as Zeeman and Göbel assert, then we have no reason to think them either discrete, continuous or anything else. Couldn't we then choose a topology whose induced topology along photon tracks is coarser, rather than finer, than the Euclidean topology—for example, one whose induced topology is the indiscrete topology. I think that with equal justification we could.[5] After all, what is important for getting the homeomorphism group to equal some symmetry group of a geometric object field is that one be able to topologically differentiate between the inside and the outside of the null cone. We are tagging, as it were, the relevant curves in the space-time as a means of cutting down the size of the symmetry group. Seen in this light the proposal to change the topology is, I think, much less interesting.

Finally, and in order to forestall some possible misunderstandings, I will outline the positive consequences of the preceding discussion. I am certainly not arguing that there could never be circumstances in which a change of topology might be considered desirable, merely that the arguments presented for the Zeeman-Göbel topologies do not constitute such a circumstance. Thus I am not in disagreement with the broad picture presented in my outline of WRT itself, at least with respect to one point: there are indeed many definable topologies for space-times, though some are more naturally related to the given structures of the space-time. I have argued that the manifold topology has this natural relation. The other definable topologies should, perhaps, just be treated instumentally. That is, certain problems may recieve an elegant formulation in topology but be cumbersome at the level of the metric, and in this circumstance employing a non-standard topology may have pragmatic justifications. But the non-standard topology is employed only for the purpose of formulating such general solutions rather than as an ubiquarian replacement for the manifold topology. I think that, as things now stand, this is the best that we can hope for: once we regard a topology as the natural, or 'real', topology for space-time then we can do no other than construe the others instrumentally, since they differ on substantive physical issues (like the proper induced topology on a light-ray). This represents my disagreement with WRT: a topology cannot be correct if it does not give the right induced topology on physical curves (or generally, submanifolds). I've argued that on these grounds there is no reason to think that a fine topology is correct. Thus trickle-down realism trickles down unequally: it favours, Zeeman and Göbel notwithstanding, the standard manifold topology for a space-time.

Dept. of Philosophy, University of Adelaide
Adelaide, South Australia

An earlier version of this paper was read at the 1983 Australasian Association of Philosophy Meeting held in Adelaide. I wish to thank all those who gave me their advice and criticisms there. A special debt of gratitude is owed to Graham Hall of the University of Aberdeen who refused to be satisfied with earlier arguments that I put to him.

---

[5] Provided of course that such a topology is indeed definable.

REFERENCES

ATIYAH, M. F. [1984]: 'The Yang-Mills equations and the structure of 4-manifolds', *Global Riemannian Geometry* (ed.) T. J. Willmore and N. J. Hitchin, Ellis Horwood, Chichester.

BISHOP, R. L. and GOLDBERG, S. I. [1980]: *Tensor Analysis on Manifolds* Dover, New York.

CEL'NIK, F. A. [1968]: 'Topological structure of space in relativistic mechanics', *Soviet Math. dokl.* **9**, 1151–1152.

CHRISTENSON, C. O. and VOXMAN, W. L. [1977]: *Aspects of Topology* Marcel Dekker, New York.

DE FRISE-CARTER, L. [1975]: 'Conformal groups and conformally equivalent isometry groups', *Commun. Math. Phys.* **40**, 273–282.

DONALDSON, S. K. [1983]: 'An application of guage theory to four dimensional topology', *J. Differential Geometry*, **18**, 279–315.

EHLERS, J., PIRANI, F. A. E. and SCHILD A. [1972]: 'The geometry of free-fall and light propogation' in *General Relativity* (ed.) L. O'Raifeartaigh, O.U.P., Oxford.

EHLERS, J. and SCHILD, A. [1972]: 'Geometry in a manifold with projective structure', *Commun. Math. Phys.* **32**, 119–146.

FREEDMAN, M. H. [1982]: 'The topology of four-dimensional manifolds', *J. Differential Geometry*, **17**, 357–453.

GÖBEL, R. [1976a]: 'Zeeman topologies on space-times of general relativity theory', *Commun. Math. Phys.* **46**, 289–307.

GÖBEL, R. [1976b]: 'The smooth-path topology for curved space-time which incorporates the conformal structure and analytic Feynman tracks', *J. Math. Phys.* **17**, 845–853.

GOMPF, R. [1983]: 'Three exotic R^4's and other anomalies', *J. Differential Geometry*, **18**, 317–328.

GUILLEMIN, V. and STERNBERG, S. [1984]: *Symplectic Techniques in Physics*, Cambridge University Press, Cambridge.

HAWKING, S. W., KING, A. R. and McCARTHY, P. J. [1976]: 'A new topology for curved space-time which incorporates the causal, differential, and conformal structures', *J. Math. Phys.* **17**, 174–181.

HU, S.-T. [1964]: *Elements of General Topology*, Holden-Day, California.

KELLEY, J. L. [1955]: *General Topology*, Springer-Verlag, New York.

KOBAYASHI, S. and NOMIZU, K. [1963]: *Foundations of Differential Geometry Vol. 1*, Interscience, New York.

MASHFORD, J. S. [1981]: 'A non-manifold theory of space-time'. *J. Math. Phys.* **22**, 1990–1993.

PENROSE, R. [1984]: 'Donaldson's Moduli Space: a 'Model' for Quantum Gravity' in *Quantum Theory of Gravity* (ed.) Stephen M. Christensen, Adam Hilger, Bristol.

STEEN, L. A. and SEEBACH, J. A. (Jr.) [1970]: *Counterexamples in Topology*, 2nd Edition, Springer-Verlag, New York.

STONE, A. H. [1948]: 'Paracompactness and product spaces', *Bull. Amer. Math. Soc.* **54**, 977–982.

ZEEMAN, E. C. [1967]: 'The topology of Minkowski space', *Topology* **6**, 161–170.

Brit. J. Phil. Sci. 39 (1988), 263–275 *Printed in Great Britain*

REVIEW ARTICLE

The Value of a Fixed Methodology*

---

1 *Introduction: Scientific Change and its Problems*
2 *The 'Reticulated Model of Change' and its Failure*
3 *The 'Hierarchical Model of Change' Revisited*
4 *Conclusion*

I INTRODUCTION: SCIENTIFIC CHANGE AND ITS PROBLEMS

Larry Laudan's latest book focusses on disagreements in science and on how such disagreements are resolved. Very briefly, Laudan holds that there is more disagreement in science then can be coped with by the accounts of scientific change given by such philosophers as the logical positivists, Reichenbach, Hempel and Popper (and by such sociologists of science as Merton); but that the fact that such disagreements are 'often' 'definitively' resolved is incompatible with the newer 'holist' or 'big picture' view associated with Kuhn and others. Laudan therefore sets out to develop his own explanatory model of scientific change which will allow *both* for the possibility of wide-ranging (and rational) disagreement *and* will explain how such disagreement might eventually be definitively resolved.

He sees the recent history of science studies in something like the following terms. The 40s and 50s were dominated by the 'consensual view'. On this view scientists standardly agree (about more or less everything and more or less all the time)—disagreements about theories *do* occasionally break out, but are quickly resolved by appeal to shared methods and goals. Then came Kuhn. His principal role was to overturn the consensual view by showing that disagreement is endemic in the scientific enterprise: scientists' disagreements about factual matters are standardly deeper and longer lasting than had generally been supposed and, more importantly, they need not be readily resolvable, because disagreement often extends beyond the factual to the *methodological* level, and even to the level of the appropriate aims and goals for science. Kuhn thus ushered in the '"new wave" preoccupation with dissensus'.

---

* Laudan, L. [1984]: *Science and Values. The Aims of Science and their Role in Scientific Debate.* University of California Press. Pp. xiv+149. US $14.95.

An explanatory 'model of consensus formation' developed alongside the consensual view: the 'hierarchical model' (which Laudan associates with such philosophers as Reichenbach, Popper and Hempel). This model identifies three levels of scientific commitment—factual, methodological and 'axiological'. It assumes that methodology governs factual matters and axiology governs methodology. Factual disagreements (understood, of course, in a broad sense which includes theoretical disagreements) are standardly resolved by appeal to shared methodological standards: if two scientists disagree about which is the better of two theories, they will quickly restore agreement by consulting their *shared* appraisal criteria which will rank the two theories for them. However, Laudan sees it as one of the important truths learned from Kuhn that disagreements between scientists can extend beyond the substantive level to the level of methodology (p. 25). Such disagreements need not embarrass the hierarchical model—since appeal can be made to the axiological level, where shared goals and aims may decide the methodological disagreement (ibid and p26). The hierarchical model *is*, however, in trouble, if disagreement extends to the top 'axiological' level, and it *does*:

> The history of science is rife with controversies between, for instance, realists and instrumentalists, reductionists and antireductionists, advocates and critics of simplicity, proponents of teleology and advocates of purely efficient causality. At bottom, all these debates have turned on divergent views about the attributes our theories should possess (and thus about the aims of scientific theorizing). (p. 42)

A rival to the hierarchical model of scientific change developed alongside the 'new wave preoccupation with dissensus': the 'big picture' or 'holist' model. This has scientists disagreeing at all three levels and indeed it sees change as occurring at all three levels simultaneously. 'Paradigm shifts' involve not only changes in theoretical commitment, they also involve changes in ideas about how to appraise theories and even in ideas about the appropriate aims for science. But then nothing is left fixed to act as neutral arbiter between two successive paradigms. This means that, while the 'holist' model can certainly accommodate widespread 'dissensus', it cannot *explain* 'how—short of sheer exhaustion or political manipulation—scientific disagreements are ever brought to closure' (p. 16). Yet Laudan sees it as a fact that such disagreements are 'often' and 'definitively' resolved.

On Laudan's account, then, the development of post-war science studies has resulted in a fundamental problem. There is too much disagreement in science for the 'hierarchical model' to handle; yet more often than not such disagreement is (at any rate *eventually*) resolved in apparently rational fashion and this is inconsistent with the 'big picture' model—the only alternative model so far available. The main aim of Laudan's book is to develop a third model which delivers the best of both worlds.

Laudan's interesting attempt fails. In section 2 I argue that Laudan's 'reticulated model' is not a genuine third alternative. And in section 3 I argue that Laudan has both seriously misidentified the older, 'hierarchical view', and seriously overestimated the impact on it of Kuhnian criticisms. My conclusion is that the 'older' approach aimed at laying down *fixed* principles of scientific theory-appraisal is the only alternative to relativism.

2 THE 'RETICULATED MODEL' AND ITS FAILURE

Laudan's reticulated model retains the three levels of its predecessors. Like the holist model, it allows for changes at all three levels, but it insists that the holist idea of *simultaneous* change at all three levels is historically quite inaccurate. Like the hierarchical model, it sees shared methodological rules as justifying certain theory-choices, but it insists that the hierarchical idea of *one-way* justification is quite wrong. Instead, accepted theories may also 'constrain' methods and accepted methods may dictate changes in presumed aims and goals. Thus Laudan's reticulated model involves

> a complex process of mutual adjustment and mutual justification going on among all three levels of scientific commitment. Justification flows upward as well as downward in the hierarchy, linking aims, methods, and factual claims. No longer should we regard any one of these levels as privileged or primary or more fundamental than the others. Axiology, methodology and factual claims are inevitably intertwined in relations of mutual dependency. (pp. 62–3).

Suppose, then, that the holist points to an allegedly 'revolutionary' change in 'world view' from one theory-methods-aims triad (T_1, M_1, A_1) to another (T_2, M_2, A_2). Assume that (T_1, M_1, A_1) was indeed accepted by the 'scientific community' at one stage and that (T_2, M_2, A_2) was accepted at a later stage. The holist sees the change as occurring all at once—thus precluding rational explanation. The reticulated modeller has at his disposal several possible 'rational paths' from one to the other. For example,[1] T_2 may come to be preferred to T_1 while M_1 and A_1 are still in force; M_2 is then proposed as possibly superior to M_1; the decision as to whether M_2 is indeed preferrable to M_1 will generally involve both A_1 and T_2 (it is a question of whether M_1 or M_2 is 'optimal for securing A_1' and this is 'typically . . . an empirical matter', which involves our best theory—in this case T_2). Assume that M_2 emerges victorious; it may later turn out 'that virtually none of the theories accepted by the scientific community as instances of good science', presumably including T_2, 'exemplify the values expressed in A_1'; and this may require (and supply the rationale for) a switch to A_2.

'Feedback', 'mutual dependency', 'upward and downward' justification all leading to the 'rationalisation' of change: Laudan's 'reticulated model' sounds just the ticket. Unfortunately it is a ticket onto the rocks.

---

[1] This is Laudan's 'tell tale' of paradigm change, sketched around pp 76–8. The same failing I point to in the text surely afflicts his other schematic cases too: a theory change cannot at the same time be explained as rational by a methodology and explain as rationed a *change* in methodology.

Assume for the moment that the methodologies involved here, that is, M_1 and M_2, are the norms which *really* govern the preferences of the scientists concerned. (As Lakatos forcefully argued such '*implicit* methodologies' might differ radically from the norms which the scientists *explicitly* articulate and defend when indulging in published methodological reflection.) Next assume with Laudan that theory T_2 is accepted over T_1 while M_1 is in force. I presume that this implies that M_1 really does—'objectively'—rank T_2 more highly than T_1. (If not then Laudan's reticulated account immediately and clearly breaks down as an *explanation* of change.) I also take it that T_1 and T_2 are supposed to be the only serious rivals around at the time. It follows that M_1 pronounces T_2 the (temporarily) best available theory.

But how, then, could T_2 possibly 'constrain' a change in M_1, as Laudan's second step requires? M_1 has done all that T_2 could ask of it: ranked it top of the available theories. If the acceptance of theory T_2 really required the switch to methodology M_2 then T_2's initial acceptance, *while the earlier methodology M_1 was in force*, could not have been rational.

The problem can be seen very sharply in Laudan's favourite historical example. He sees the early 19th Century revolution in optics as having involved radical changes not only at the theoretical, but also at the methodological, level. Not only the Newtonian theory of light, but also Newtonian inductivist methodology, was predominant in the 18th Century. This methodology—on Laudan's account—banned all genuinely theoretical entities and put no premium on *predictive* success. The revolution saw not only the acceptance of the wave theory of light but also the acceptance of a more liberal 'hypothetico-deductive' methodology which sanctioned theoretical entities and stressed the probative value of novel prediction. In earlier accounts Laudan was quite explicit: the methodological change was *real* not just *professed*, and it was brought about by the theoretical change—a theory which gives a central role to the space-filling but invisible and intangible 'luminiferous ether' could hardly co-habit for long with a methodology which anathematised theoretical entities.

But, aside from any question of the historical accuracy of this account, it faces a clear (and devastating) logical problem: if Newtonian inductivism really were in force at the time Fresnel developed his theory (and *if* Newtonian inductivism really does ban genuinely theoretical entities) then the acceptance of that theory by the scientific community could not have been rational. Conversely, of course, if the initial acceptance of the wave theory was rational, then Newtonian inductivism (as described by Laudan) was not really in force at the time. In neither case can there have been a real shift in methodology which can be explained as the rational response to the prior acceptance of the wave theory.

Some remarks suggest that Laudan may now have moved towards an account of this episode more in terms of 'explicit methodology'. It is of course perfectly possible to construct a logically flawless account of this episode in these terms. This would go as follows. Fresnel's theory was *intuitively* accepted as enormously successful—as really superior to the Newtonian corpuscular theory on scientists' real, *implicit* standards. However, scientists in the 18th and early 19th Centuries had been spoilt by the success of Newtonian physics, were under the *illusion* that that success had been achieved by applying the strictly positivistic Newtonian method and hence tended, in their *explicit* methodological writings, to emphasise induction and 'deduction from the phenomena' as the correct scientific procedures. Their intuitively high appraisal of the wave theory, coupled with the glaringly obvious fact that the 'luminiferous ether' is a highly theoretical entity, helped free them from this (explicit) methodological illusion, and helped them to see that 'deduction from the phenomena' is a myth (that is, a myth at the level of *real*, 'implicit' methodology). On this *second* account, the 'methodological revolution' involved no *real* change in methodology, but consisted in scientists' explicit methodological pronouncements coming much more closely into line with the implicit methodological standards they had in fact always applied.

The problem with this second account from Laudan's point of view is that it provides no example of 'reticulation'. On this account there was no change at the real, 'implicit' methodological level. Indeed it is the (presumed) fact that Fresnel's theory was superior to its predecessor according to *unchanging* implicit methodological standards which explains the shift to that theory as rational.

Laudan is in a dilemma. If the methodological change that he claims has occurred is change in implicit methodology, then his 'reticulated' account supplies no rationale for scientific change. In this case in fact his 'reticulated model' collapses into 'big-picture relativism'. If, on the other hand, those methodological changes are only in 'explicit' methodology, then they present no challenge to the 'hierarchical model' (which, at any rate as I understand it, is concerned solely with real, implicit methodology). In neither case does the 'reticulated model' stand as a genuine third alternative.

This conclusion is cold comfort if the 'hierarchical model' of change is in as serious trouble as Laudan alleges. For that would leave 'big-picture relativism' as the only alternative—and an abhorrent alternative at that. Fortunately, as I shall argue in the next section, Laudan both misdescribes the 'hierarchical view' and greatly exaggerates the impact on it of Kuhnian criticisms.

3 THE 'HIERARCHICAL MODEL' REVISITED

Laudan takes it that Kuhn and others have demonstrated that there have been important changes at the methodological level in science, and at the level of aims and goals. These changes cannot be accommodated within the 'hierarchical' view. But if we are to decide whether or not there has been methodological change in the history of science, we clearly need first to decide exactly what counts as a methodological principle. And ordinary usage is far from unambiguous on this point. It seems to me that Kuhn (and Laudan) take a *much wider* understanding of methodology than the earlier 'positivist' philosophers: to the extent that *none* of Kuhn's and Laudan's examples of alleged methodological change are changes in methodology as understood by those earlier people. This means that, despite the rhetoric, this older view remains, in this respect at least, totally unaffected by these Kuhnian criticisms. Since it need not accept that there has been any methodological change *in its sense*, the hierarchical view need not bring in aims and goals to arbitrate such changes, as Laudan suggests it will. This is fortunate since the 'axiological' level would be quite unequal to this task. Let me expand on these compressed remarks.

(i) *The 'older view' of methodology*

Just as scientists do not usually indulge in meta-science but instead get on with their scientific work, so methodologists do not usually indulge in meta-methodology. Nonetheless I take it that Carnap, Hempel, Reichenbach as well as earlier thinkers such as the French conventionalists Duhem and Poincaré, all thought of themselves as aiming to articulate the principles which *invariably* govern the appraisal of scientific theories. The idea was to do for the 'logic of science' what Boole, Frege, Russell and the mathematical logicians had done for *deductive* logic: just as the latter had articulated the general and unchanging principles of valid inference, so those methodologists were trying to find the general and unchanging principles of 'inductive logic', and in particular of when one theory was, in view of all the evidence, better supported than another. Of course, just as inferences are sometimes made which are invalid, scientific appraisals may sometimes be made which are incorrect. It is even possible that whole historical epochs may be characterised by consistently 'incorrect' appraisals; and it is certainly possible that attempts to *articulate* these principles may go badly astray. But just as the existence of invalid arguments does not challenge the abstract idea of deductive validity and certainly does not imply that the objective standards of validity have changed over time, so the existence of incorrect scientific appraisals does not challenge the idea that—at the abstract, logical level—there are fixed principles of theory appraisal, and it certainly does not imply that the real standards of appraisal have changed over time.

In order to play the envisaged role, the principles of scientific theory appraisal were clearly going to need to be principles of great generality and abstractness, relying on no controversial 'metaphysical' assumptions about the structure of the universe. There were of course (and continue to be) hotly debated differences over the correct exact articulation of these general rules. But I take it that there was no suggestion that two different sets of principles might be the right ones for different sets of scientists at different times. Indeed there was a great measure of agreement between the different philosophers *at the intuitive level*—agreement that the crucial question is how well the different theories stand up to the empirical evidence. The differences tended to be over such further questions as whether or not the shared intuitive judgments about empirical support were captured by some formal, probabilistic system. On this 'older' conception, then, even the best scientists might, on occasion, make a methodological mistake, and the sorts of 'explicit' methodological pronouncements that were fashionable might certainly change over time. But the standards of correct scientific appraisal of theories are assumed to be fixed and unchanging (and therefore highly abstract and general); and it is the task of the philosopher of science to articulate these unchanging standards sharply and clearly. (One fairly recent advance has been the recognition that the philosopher of science is not likely to be able to do this by *a priori* cogitation but must instead make a detailed study of the progress of science. This does not however entail that the standards he is aiming to articulate are themselves historical, that is, subject to change.)

(ii) *These 'older' philosophers did not hold that 'axiology' governs methodology*

Laudan explicitly cites Carnap, Hempel, Reichenbach and Popper as proponents of the 'hierarchical model'. But as I understand these philosophers, they never really envisaged genuine methodological disagreements which would require arbitration at some other level; *a fortiori* they did not envisage such disagreements being resolved *via* discussion of the aims and goals of science. If two people disagree over whether or not an inference is deductively valid (and assuming this disagreement does not depend on linguistic vagueness) then one of them must be *wrong*. Similarly if two scientists disagree over which of two theories is presently better supported by the evidence, then, on the 'older' view in philosophy of science, unless this is a 'grey case', one of the scientists must be wrong. In other words, the 'older' view is *not* hierarchical in Laudan's sense: methodology governs factual disagreements alright, but there are no real, implicit methodological disagreements for 'axiology' to govern.

This seems to me fortunate since a discussion of the aims and goals of science would surely be quite unsuited to settling methodological disputes. Views about aims and goals seem altogether more ephemeral, more 'philosophical' than judgements about which scientific theory is presently best supported empirically. Indeed, and contrary to what Laudan more than once asserts, philosophers with very different ideas about the aim of theoretical science will standardly agree about which theory is currently the best available (though their *explicit* reasons for regarding that theory as best will often differ). So for example an instrumentalist and a realist will standardly agree about which theory is currently the best available. But one will claim that that best theory is our best attempted description of reality, while the other claims that the same theory is the best, most comprehensive and efficient codification of the phenomena. They will also disagree over how that best theory is to be interpreted—does it really say that there are subatomic particles or that there is a real, curved structure of spacetime; or is that all 'as if' talk? But nonetheless (syntactically) the same theories will appear in the textbooks written by the instrumentalist and by the realist.

(iii) *Kuhn (and those influenced by him) adopt a very different, because much broader, conception of methodology*

The old idea that Newton's theory was better than Aristotle's for the same general reason that Einstein's theory is better than Newton's has not, so far as I am concerned, been discredited by recent studies of science. It is true that philosophers have not yet agreed on exactly what the 'same general reason' is, and they have certainly not agreed on what exactly follows from the judgement that one theory is better than another. Nonetheless, despite the fact that their explicit methodological remarks will be very different, Duhem, Poincaré, Carnap, Reichenbach, Popper and others would all surely produce the same ranking: Einstein, Newton, Aristotle. This may not be very much but it is still, I believe, the key to scientific rationality.

Why is it so widely believed that Kuhn and Kuhn-influenced studies of science have destroyed this old idea? The answer, as I already indicated, is that Kuhn operates with a much broader conception of methodology. Of course, 'methodology' is a vague term, and many more principles than these very general abstract ones involved in ranking theories in the light of the evidence can perfectly properly be included under it without straining ordinary usage.

One broad class of principles which are certainly 'methodological' in a sense, but are not the sort of thing the 'positivists' had in mind arises as follows. An important feature of science which *did* tend to drop out of sight in earlier 'positivist' accounts is the multi-level nature of science's substantive, factual components. At any stage—especially in a well-developed science like physics—there will be a whole range of accepted theories, extending from observational and auxiliary theories (the theory of how light affects photographic emulsion, say) through specific theories (Fresnel's wave theory of diffraction, say) to more general theories (the general wave theory of light, say, that light is *some sort* of disturbance in *some sort* of mechanical medium) and on to *very* general claims of a metaphysical character (light is some sort of mechanical and deterministic phenomenon). The relationships between theories at different levels are important. For example, the levels supply at any rate a rough pecking order in case of empirical refutation. Once a general idea has proved its mettle, the standard reaction to the empirical refutation of a specific theory will be to look for a new specific theory which embodies the same general idea. Replacements for the general idea will standardly be sought only once a *series* of specific theories built around them have proved unsuccessful. For example, Fresnel's 1819 theory of light was refuted by the failure of oppositely polarised light beams to produce interference fringes in appropriate circumstances. His reaction was to hold on to the *general* wave theory of light and in fact to use the initially refuting experimental results to determine a new specific theory within that same general framework. It was only a series of failures by Maxwell and others to construct empirically adequate 'mechanistic reductions' of Maxwell's more general electromagnetic theory which forced the eventual abandonment of the idea of a mechanical medium. And even then the still more general principles, like determinism, continued to operate. Until even that principle was eventually surrendered (at any rate by most scientists).

While these more general principles—either 'scientific' like the general wave theory or 'metaphysical' like the principle of determinism—are accepted by science, they play a *dual* role: both substantive *and* heuristic. Not only are they accepted parts of scientific theory, making assertions about the world ('light is a wave motion in some medium'); they also operate as heuristic principles ('if your initial specific theory fails, look for another specific theory which still has the general features of making light a wave motion'). It by no means stretches normal usage, then, to regard these principles in their heuristic role as part of the then accepted 'methodology of science'. But if these principles *are* categorised as methodological, then of course it is no news that accepted 'methodology' has changed along with (though more slowly than) substantive science, nor that substantive, theoretical considerations have played important roles in 'methodological' change. Moreover, at any one stage in the development of science there may be 'methodological disagreements' between particular major scientists. Einstein's famous resistance to the quantum theory was over the question of whether or not the time had come to abandon determinism and the idea that all dynamical variables must *always* have well-defined values as general requirements on all physical theories. It is vital to note, however, that Einstein never for a moment questioned the present superiority of the quantum theory over all rivals—in the very abstract terms deemed methodological by positivists and others. Einstein clearly regarded the quantum theory as evidentially the best supported theory then available. It's just that he did not believe that it could be true, and hence had different views than most of his contemporaries about how that part of science should progress. Similarly, David Brewster—the most famous (and perhaps only competent) British resister of Fresnel's wave theory—readily accepted the superior empirical power of the wave theory as compared with its rivals. Nonetheless, he felt that the way forward—that is, the way to produce a new theory still better empirically supported than Fresnel's wave theory—lay outside of the general wave approach.

None of these changes in, and disagreements about, 'methodology' in this wider sense, at all challenges the views of Carnap, Hempel *et al.*, all of whom had the much narrower construal of 'methodology' in mind. Indeed these 'older' philosophers would *explain* the disentrenchment of the general wave approach in optics or the disentrenchment of the still more general principle of determinism, precisely on the grounds that scientists eventually found themselves unable, despite repeated efforts, to produce specific theories which entailed the general wave theory, or which entailed determinism, *and which maximally satisfied the unchanging, abstract formal principles of good science.*

(iv) *What counts as a methodological principle for Laudan?*

Laudan is never entirely clear about what is to count as a methodological principle. As we saw, he takes it as already established (principally by Kuhn) that there have been important changes at the methodological level in science, and plunges straight into the attempt to produce a rational model for such changes. This, I think, implicitly commits him to the 'wider view' of methodology, which includes the general descriptive/heuristic principles of the type just described. As we again saw earlier, he also seems to count as a methodological change any change in the kind of *explicit* methodological pronouncements scientists are likely to produce at any historical epoch. However his chief explicitly cited examples are of still different kinds of case.

First, as I already remarked, he claims that an important methodological change occurred in tandem with the early 19th Century revolution in optics—a shift which allegedly lifted a ban on theoretical entities and which led, for the first time, to a premium on the prediction of temporally novel facts.

This is wrong for a variety of reasons. The pre-19th Century 'ban' on theoretical entities could hardly have been part of real, implicit methodology since the universal force of gravity was certainly an accepted part of physics before the early 19th Century, and is an especially clear example of a theoretical entity. Secondly, successful prediction has surely *always* been a hallmark of a good scientific theory. Here, for example is Leibniz writing 250 years before Fresnel:

> [It is] the greatest commendation of a hypothesis (next to truth) if by its help predictions can be made even about phenomena not yet tried. (Letter to Conring, 1678)

Laudan has, I think, been misled by a certain ambiguity about 'prediction'. As I have tried to explain elsewhere (see my [1985] and [1988]) successful predictions need not be of *temporally novel* facts. The methodologically important distinction is between those empirical results which a theory 'predicts', that is, *yields naturally without any prior contrivance*, and those results for some disease, a plausible alternative is always that any observed improvement was due to placebo-effects rather than the specific action of the drug. Hence the need to perform tests 'double blind'—that is, to test the specific so long as these results 'fall out of' the theory. Here, for example, is the author of an excellent recent textbook on Newtonian mechanics:

> like any other good theory in physics, [Newton's theory of universal gravitation] had predictive value; that is, it could be applied to situations besides the ones from which it was deduced. Investigating the predictions of a theory may involve looking for hitherto unsuspected phenomena, or it may involve recognizing that an already familiar phenomenon must fit into the new framework. In either case the theory is subjected to searching tests . . . With Newton's theory of gravitation, the initial tests resided almost entirely in the analysis of known effects—but what a list! (French [1971], pp. 5–6)

This, I think, supplies the clue to the real shift Laudan has misidentified. Prior to the 19th Century the outstandingly successful scientific theory was Newton's gravitational theory. It happened that all the successful predictions in the early career of this theory were of already known effects, such as 'precession of the equinoxes'. (The successful 'novel' predictions—such as the existence of Neptune—were of course 19th Century phenomena.) In the early 19th Century another outstandingly successful theory appeared—Fresnel's wave theory of light. This had fairly immediate 'novel' predictive success—with for example the 'white spot' affair, and less ambiguously, with Hamilton's prediction of conical refraction. This put the methodological spotlight to some extent on novel predictions and occasionally confused even astute methodologists like Whewell into holding the view that there is something special about predictive novelty. But the 'something special' is, at best, psychological—and, as I have shown in my [1988], the best contemporary scientists were just as impressed by Fresnel's theory's 'prediction' of the *already known* details of straightedge diffraction as with any of its other—temporally novel—successes. The shift, then, was simply that in the 19th Century there were genuinely novel predictive successes for methodologists to become confused about.

The second example which Laudan cites as an example of a methodological change is the methodological 'discovery' of the double-blind principle for clinical trials. Clearly the double-blind principle would be treated in medical textbooks as part of the *methodology* of clinical trials—but clinical trials were not always performed double-blind. Here then is surely a clear-cut example of a methodological innovation. However, the 'positivist' could equally well tell the story of the invention of the double-blind 'methodology' while retaining his view that methodology *in his sense* remained invariant. He would surely say that it is an invariable principle of good scientific practice that, whenever possible, theories should be tested *against* plausible rival theories. The *substantive* (non-methodological) discovery of placebo-effects entailed that, for any theory which claims that a particular drug, say, has some specific efficacy which have merely been assimilated into the theory, that is, those which have been *used in the construction of the theory*. Although there is some confusion over usage, scientists indeed often talk of predictions of already known results, theory *against* the alternative placebo hypothesis. In other words, the old-fashioned position would see the innovation here as the result of plugging a new substantive discovery (that placebo effects sometimes operate) into an *invariant* methodological principle (theories should, whenever possible, be tested against plausible rivals).

In sum, then, I see nothing in Kuhn or elsewhere to challenge the view that methodology in the narrower sense has changed with the development of science. This is fortunate since, without invariant principles of good science, the whole idea of explaining the development of science as a *rational* process has surely been abandoned.

4 CONCLUSION

Assuming that changes in real, 'implicit' methodology are at issue, then Laudan's 'reticulated model', as it stands, collapses into relativism. If no principles of evaluation stay fixed, then there is no 'objective viewpoint' from which we can show that progress has occurred and we can say only that progress has occurred *relative to the standards that we happen to accept now*. However this may be dressed up, it is relativism. Without fixed standards, no amount of 'mutual adjustment . . . among all three levels of scientific commitment' can avoid it.

As he usually presents it, Laudan's 'reticulated model' presupposes no constraints on the methodological changes that *may* take place. The model *could* be saved from collapse into relativism by adding the requirement that there be an invariant core M* which is not subject to change and hence which is common to all the different methodologies M_i. Although the whole thrust of Laudan's presentation is that the 'reticulated model' is to be a definite *rival* to, indeed replacement for, the 'hierarchical' one, some remarks do suggest that he might be ready to concede that some 'core' methodological principles do indeed remain, and must remain, invariant. But then the 'reticulated model' becomes merely an *elaboration of*, rather than a rival to, 'hierarchism'. The defender of the 'older' view will, of course, readily concede that there have been methodological changes in the wider sense of 'methodology', but will insist that these very changes have themselves been adjudicated by the principles of methodology *in his narrower sense*. It is thus these principles in the core M* which are the key to the rationality of scientific progress.

If Laudan would allow that there is an invariant core M* to all his changing methodologies M_i and insist merely that his 'reticulated model' supplies a congenial framework within which to describe changes in methodology in the wider sense, then there would be no objection to it. Except that his presentation of it as a genuine third alternative totally obscures the real situation.

There is an important decision to be made: either there is an invariant core M* of methodological principles or everything is open to change. *With* such an M*, the 'reticulated model' is an elaboration, or notational variation, of the older 'hierarchical' view; *without* such an M* the model collapses into relativism. Either way, Laudan's 'third' is excluded as an independent, rival position.

JOHN WORRALL
London School of Economics

REFERENCES

FRENCH, A. P. [1971]: *Newtonian Mechanics*. London: Nelson.

WORRALL, J. [1985]: 'Scientific Discovery and Theory-Confirmation' in J. Pitt (ed.): *Change and Progress in Science*. Dordrecht: Reidel.

WORRALL, J. [1988]: 'Fresnel, Poisson and the White Spot: the Role of Successful Predictions in the Acceptance of Scientific Theories' in Gooding, Schaffer, and Pinch (eds.): *The Uses of Experiment*. Cambridge University Press.

The British Society for the Philosophy of Science

Annual Conference 1988
University of Hull: September 23–25 1988

Professor Ian Hacking (University of Toronto) *Human Kinds*

Dr Jeremy Butterfield (University of Cambridge) *Peaceful Coexistence in Microphysics: How's Chances? (Quantum Mechanics)*
Dr Harvey Brown (University of Oxford) *A Reply*
Chair: Dr Peter Gibbins (Open University)

Professor Susan Haack (University of Warwick) *What is 'the Problem of the Empirical Basis' and does Johnny Wideawake solve it?*
Professor: John Watkins (L.S.E.) *A Reply*
Chair: Professor Hugh Mellor (University of Cambridge)

Professor Anthony O'Hear (University of Bradford) *Evolution and Knowledge*
Mr Michael Smithurst (University of Southampton) *The Elusiveness of Human Nature*
Chair: Mr Paul Gilbert (University of Hull)

Professor John Pheby (Birmingham Polytechnic) *The Relevance of Laudan for Economists*
Dr Sheila Dow (University of Stirling) *Keynesean Thought—An Alternative to Dualism*
Chair: Dr Donald Gillies (University of London)

Papers by Research Students

Further details and application forms may be obtained from the Conference Organiser:

Dr A. J. Dale,
Department of Philosophy,
University of Hull,
Hull, U.K.

Physical Interpretations
of
Relativity Theory

LONDON: SEPTEMBER, 1988

The British Society for the Philosophy of Science is sponsoring an international conference, of three days' duration, "PHYSICAL INTERPRETATIONS OF RELATIVITY THEORY", to review the development, status, and potential of the various physical interpretations of the Relativistic Formal Structure. It is planned to open on Friday, 16th September, and to close on Monday, 19th September, 1988. The location will be Imperial College, London.

SECTIONS INCLUDE:

"Physical Time and Relativity"
"Privileged Reference Systems in Modern Physics and Cosmology"
"Relativity and the Nature of the Physical Vacuum"
"Ether Theory in the late 20th C."
"Relativity, Cosmology and Physical Theories of Related Phenomena"
"Experimental Aspects of Relativity"
"Formal Structures, Physical Interpretations and the Philosophy of Modern Physics"

The organising committee includes Prof. C. W. Kilmister (London); Dr. P. Kroes, (Eindhoven); Dr. T. Sjodin, (Brussels); S. J. Prokhovnik, (New South Wales); Dr. S. V. Clube, (Oxford); Dr. M. F. Podlaha, (Munich); Dr. T. Morris, (N.P.L. Teddington); Dr. J. Whealton, (Oak Ridge, USA); Prof. G. Spinelli, (Milan); Prof. S. Bergia, (Bologna); and M. C. Duffy, (Sunderland).

For more information, contact
Conference Co-ordinator: M. C. Duffy,
Mechanical Engineering Dept., Sunderland Polytechnic,
Chester Road, Sunderland SRL 3SD, U.K.
Tel: 091 567 6191 Ext. 107

Philosophy and Cosmology

On September 21–25, 1988 an international conference will be held at Colorado State University on

'The Origin of the Universe: Philosophical
and Scientific Perspectives'

This is the second annual conference sponsored by the Endowment for Applied Philosophy and the Department of Philosophy.

Invited speakers include:

George Gale
Milton Munitz
David Schramm
John Leslie
Donald Page
Frank Tipler
Charles Misner
John Polkinghorne
Anthony Zee

For further information contact:
Richard Kitchener,
Department of Philosophy,
Colarado State University,
Fort Collins,
CO 80523,
USA.
(303) 491-6315.

Journal of Philosophical Logic

Editor-in-Chief

Richmond H. Thomason, *Dept. of Philosophy, University of Pittsburgh, USA*

The *Journal of Philosophical Logic* is the only journal specializing in philosophical logic and utilizes formal methods of dealing with topics in logical theory.

Subscription Information **ISSN** 0022–3611
1988, Volume 17 (4 issues)
Institutional rate: Dfl. 228.00/US$108.00 incl. postage/handling
Private rate: Dfl. 82.00/US$31.50 incl. postage/handling
Special rate for ASL members: Dfl. 62.00/US$21.50 incl. postage/handling

Private subscriptions should be sent direct to the publishers

Kluwer Academic Publishers
P.O. Box 989, 3300 AZ Dordrecht, The Netherlands
101 Philip Drive, Norwell, MA 02061, U.S.A.

GRAZER
PHILOSOPHISCHE STUDIEN

BAND 31 1988 VOLUME 31

Aufsätze *Articles*

Wozu könnte Moral gut sein?
oder Kant, das Gefangenendilemma und die Klugheit . . . . Rainer HEGSELMANN
„Wie du mir, so ich dir“ statt „Was du nicht willst, das man dir tu’,
das füg’ auch keinem andern zu“ –
Eine (partielle) Rehabilitation Hammurabis . . . . Hans LOTTENBACH
France Veber’s Theorie of Value . . . . Seppo SAJAMA
Kontextualität und Transkontextualität . . . . H. R. FISCHER, F. B. SIMON
Concepts, Ideas and Definitions in Schlick’s
Allgemeine Erkenntnislehre . . . . Marie Luise SCHUBERT-KALSI
Wahrheitsdefinition und Wahrheitskriterium . . . . Geo SIEGWART
Semantische Vagheiten im Lichte der dreiwertigen Logik,
der Superbewertung und der unscharfen Logik . . . . Wolfgang G. STOCK
Direct Realism, Skepticism and Truth . . . . John PETERSON
The Existential Basis of Propositions, States of Affairs and Properties . . . . Thomas R. GRIMES
An Ontological Argument for Modal Realism . . . . Michael LOSONSKY
Braving the Perils of an Uneventful World . . . . T. HORGAN, M. TYE
Plantinga, Foundationalism and the Charge of Selfreferential Incoherence . . . . . . John GRECO
The Tradition in Retreat . . . . Kai NIELSEN
Ist Gilbert Ryle erledigt? . . . . Matthias KAUFMANN

Herausgeber *Editor*

Prof. Dr. Rudolf Haller, Institut für Philosophie, Universität Graz, Heinrichstraße 26, A-8010 Graz, Österreich/Austria.

Erkenntnis
An International Journal of Analytic Philosophy

Editors
Carl G. Hempel, *Princeton University, USA;* **Wolfgang Stegmüller,** *Universität München, FRG,* **Wilhelm K. Essler,** *Universität Frankfurt, FRG*

Erkenntnis is a philosophical journal publishing papers on foundational studies and scientific methodology covering: 'Philosophy of Science' and 'Analytic Philosophy'; philosophy of language, of logic, and of mathematics; foundational problems of physics and of other natural sciences; foundations of normative disciplines such as ethics, philosophy of law and of aesthetics; methodology of the social sciences and the humanities; history of scientific method. *Erkenntnis* was first published in 1930, and succeeded *Annalen der Philosophie*, which appeared from 1919–29.

Subscription Information **ISSN** 0165–0106
1988, Volume 28–29 (6 issues)
Institutional rate: Dfl. 452.00/US$214.00 incl. postage/handling
Private rate: Dfl. 172.00/US$72.00 incl. postage/handling

Kluwer Academic Publishers
P.O. Box 989, 3300 AZ Dordrecht, The Netherlands
101 Philip Drive, Norwell, MA 02061, U.S.A.

Linguistics and Philosophy

An International Journal

Managing Editors

David R. Dowty, *Dept. Of Linguistics, Ohio State University, Ohio, USA* and **Francis J. Pelletier,** *Dept. of Philosophy, University of Alberta, Edmonton, Canada*

Linguistics and Philosophy is a journal for studies of natural languages and will interest philosophers as well as linguists. Contributions come in the form of articles, review articles, notes, discussions, and remarks and replies.

Subscription Information **ISSN** 0165–0157
1988, Volume 11 (4 issues)
Institutional rate: Dfl. 221.00/US $105.00 incl. postage/handling
Private rate: Dfl. 86.00/US $31.00 incl. postage/handling
Special rate for students: Dfl. 80.00/US$28.50 incl. postage/handling

Private subscriptions should be sent direct to the publishers

Kluwer Academic Publishers
P.O. Box 989, 3300 AZ Dordrecht, The Netherlands
101 Philip Drive, Norwell, MA 02061, U.S.A.

INQUIRY

An Interdisciplinary Journal of Philosophy

Editor: Alastair Hannay

Vol. 31 No. 1, March 1988

- *Peter Skagestad:* On History's Witness Stand: Rubashov, Bukharin, and the Logic of Totalitarianism
- *Paul Seabright:* Objectivity, Disagreement, and Projectibility
- *Harold I. Brown:* Normative Epistemology and Naturalized Epistemology
- *B.J. Garrett:* Best-candidate Theories and Identity: Reply to Brennan
- *Andrew Brennan:* Reply to Garrett
- *Richard C. Jennings:* Alternative Mathematics and the Strong Programme: Reply to Triplett
- *Maurice A. Finocchiaro:* Science and Society in Newton and Marx (Georgy Markus, *Language and Production: A Critique of the Paradigms,* Gideon Freudenthal, *Atom and Individual in the Age of Newton: On the Genesis of the Mechanistic World View)*

- ✂

To: Norwegian University Press
P O Box 2959 Tøyen, 0608 Oslo 6, Norway

☐ Please enter my subscription to INQUIRY (4 issues per year)

Rates 1988 (postage included). **Airmailed to subscribers in the Americas**

Nordic countries only:	☐ Institutions NOK 400,-	☐ Individuals NOK 205,-
All other countries:	☐ Institutions USD 67.00	☐ Individuals USD 34.00

(Single issues: NOK 125,-/USD 21.00 plus postage)

Name: ....

Address: ....

....

Individuals must order direct from publisher

☐ Cheque enclosed ☐ Please send invoice

1-88/108A